

STIC-Biotech/ChemLib

From: Swope, Sheridan
Sent: Thursday, June 09, 2005 4:48 PM
To: STIC-Biotech/ChemLib
Subject: 10/804,457

For 10/804,457, pls search and interference search:

SID 2-3 against the NT data bases

THANKS

Sheridan Swope, Ph.D.
Patent Examiner, AU 1652
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E02C70 Remsen Bld (Mailbox)

Toby Port

STIC
10/804,457

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2005, 13:09:11 ; Search time 239 Seconds
(without alignments)
3361.556 Million cell updates/sec

Title: US-10-804-457-2
Perfect score: 2628
Sequence: 1 MKPRARGWRGLAALWMLLAQ.....SFCRQYLHLFRLSRVYKMPAL 491

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlp
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10804457/runat.13062005.173510.3511/app_query.fasta.1.647
-DB=Issued Patents NA -QPMI=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOFCU=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10804457 @CGN 1 1 105 @runat.13062005.173510.3511 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT_DSPELCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUTS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2628	100.0	1476	US-09-930-872-1	Sequence 1, Appli
2	2628	100.0	1476	US-10-217-774-1	Sequence 1, Appli
3	2392	98.6	3675	US-09-930-872-3	Sequence 3, Appli
4	2592	98.6	3675	US-10-217-774-3	Sequence 3, Appli
5	2592	98.6	4042	US-09-930-872-5	Sequence 5, Appli
6	2592	98.6	4042	US-10-217-774-5	Sequence 5, Appli
7	720.5	27.4	1407	US-09-963-791-5	Sequence 5, Appli
8	716.5	27.3	1524	US-09-963-791-9	Sequence 9, Appli
9	716.5	27.3	1770	US-09-963-791-11	Sequence 11, Appli
10	716.5	27.3	2727	US-09-963-791-1	Sequence 1, Appli
11	715	27.2	3218	US-09-369-364A-6	Sequence 6, Appli
12	706	26.9	2848	US-09-369-364A-4	Sequence 4, Appli

c	13	704.5	26.8	5357	3	US-09-392-184-5	Sequence 5, Appli
	14	695.5	26.5	3766	4	US-09-981-953A-1	Sequence 1, Appli
	15	691.5	26.3	3571	4	US-09-799-451-411	Sequence 411, App
	16	690.5	26.3	3377	4	US-09-981-953A-3	Sequence 3, Appli
	17	672.5	25.6	3160	4	US-09-963-791-25	Sequence 25, Appli
	18	598	22.8	954	4	US-09-963-791-15	Sequence 15, Appli
	19	594	22.6	1071	4	US-09-963-791-19	Sequence 19, Appli
	20	594	22.6	1317	4	US-09-963-791-21	Sequence 21, Appli
	21	594	22.6	2274	4	US-09-963-791-23	Sequence 23, Appli
	22	575	21.9	2023	3	US-09-491-522-2	Sequence 6, Appli
	23	575	21.9	2450	3	US-09-491-522-2	Sequence 2, Appli
	24	575	21.9	6692	3	US-09-491-522-1	Sequence 1, Appli
	25	571.5	21.7	2450	3	US-09-491-522-9	Sequence 9, Appli
	26	571.5	21.7	4580	3	US-09-491-522-8	Sequence 8, Appli
	27	569.5	21.7	1701	4	US-09-949-016-2634	Sequence 2634, Ap
	28	569.5	21.6	3636	4	US-09-949-016-5530	Sequence 5530, Ap
	29	568.5	21.6	1701	4	US-09-949-016-1139	Sequence 1139, Ap
	30	551	21.0	3885	3	US-09-369-364A-16	Sequence 16, Appli
	31	535	20.4	5804	3	US-09-369-364A-12	Sequence 12, Appli
	32	517.5	19.7	2625	3	US-09-369-364A-14	Sequence 14, Appli
	33	494.5	18.8	4676	3	US-09-130-491-1	Sequence 1, Appli
	34	491.5	18.7	3889	4	US-09-568-559-1	Sequence 1, Appli
	35	481.5	18.3	4858	3	US-09-392-184-1	Sequence 1, Appli
	36	450.5	17.1	3250	3	US-09-122-126B-14	Sequence 14, Appli
	37	450.5	17.1	3250	4	US-09-634-286A-14	Sequence 14, Appli
	38	450.5	17.1	3250	4	US-10-247-685-14	Sequence 14, Appli
	39	442	16.8	6659	4	US-09-321-987B-1	Sequence 1, Appli
	40	441	16.8	2853	4	US-10-009-332-2	Sequence 2, Appli
	41	431	16.4	1061	4	US-09-270-767-995	Sequence 995, App
	42	431	16.4	1061	4	US-09-270-767-16277	Sequence 16277, A
	43	427	16.2	3002	3	US-09-369-364A-1	Sequence 1, Appli
	44	399	15.2	3126	3	US-09-392-184-7	Sequence 7, Appli
	45	398.5	15.2	3638	3	US-09-369-364A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-930-872-1
; Sequence 1, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Hilbule, Erni
; APPLICANT: Carl Johan
; TITLE OF INVENTION: No. 6448388el Human Proteases and Polynucleotides Encoding the San
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/09/930, 872
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1476
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-930-872-1

Alignment Scores:
Pred. No.: 4.48e-238 Length: 1476
Score: 2628.00 Matches: 491
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-804-457-2 (1-491) x US-09-930-872-1 (1-1476)

Qy	1	MetLysProArgGAlaArgGlyTTPArgGlyLeuAlaLeuTTPMetLeuAlaGln 20
Db	1	ATGAAGCCCGCCGCCGCGATGGCGGCGCTGTGGATGCTGTGGCGAG 60
Qy	21	ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAalaProGlySerPro 40

Db 241 CACCATCAGCGCGGAGAGAGAGCGTGGCCGTTCCGAGGTGAGTCTCTTCCACCTTCGG 300
QY LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerLeuValAla 120
Db 301 CTGAAAGGCTCCAGGACGACTTCCACGTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360
QY ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 361 CCTGGCTTTATTTGTGCAGACGTTGGGAAAGACAGGCACCTAAGTCTGTGCAGACTTTACCG 420
QY 141 ProGluAspPheCysPheTyrcGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db 421 CCAGAGGACTTCTGTTCATCAAGGCTCTTGGCGATCAACAGAAATCCCTCAGTGGCC 480
QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrcPheLeu 180
Db 481 CTTTCACCTGCAGGCTTGTTCAGGCATGATACGACAGAGAGCGCATTTACTTCTTA 540
QY 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaGlnGlySerSerPro 200
Db 541 AGGCCACTTCTTCCACCTCTCATGGAACCTCGGCAGAGCTGCCAAGGCAGCTCGCCA 600
QY 201 SerHisValLeuTyrcLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
Db 601 TCCACGTACTGTACAAGAGATCCAGAGGCCCATGCTCTCTGGGGCCAGTGAGGTCTGTG 660
QY 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db 661 GTACCTCAGGACATGGAGCTGGGCATCATCAACCCCTGCACAGCAGCCTTTCGCTGTG 720
QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrcMetProGlnProPro 260
Db 721 GGACTGCCACAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCTCC 780
QY 261 LysGluAspLeuPheIleLeuProAspGluTyrcLysSerCysLeuArgHisLysArgSer 280
Db 781 AAGGAAGACCTCTTCATCTTGCAGATGATGATATAGTCTTGTTCGGGCATTAAGCGCTCT 840
QY 281 LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValAspLys 300
Db 841 CTTCTGAGTCCCATAGAAATGAGNACTGAACGTGGAGACCTTGGTGGTGTGCACAAA 900
QY 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrThrTyrcValLeuThrIleLeuAsn 320
Db 901 AAGATGATGCAAAACCATGSCCATGAAATATACCCACCTAGCTGTCTCAGCATACTCAAC 960
QY 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAsnIleAlaIleVal 340
Db 961 ATGGTATCTGCTTTATTTCAAAGATGGAACAATAGGAGGAAACATCAACATTGCAATTGTA 1020
QY 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis 360
Db 1021 GGTCTGATTCTTCTAGAAATGAACAGCCAGGACTGGTGTAAAGTCAACCCAGCACAC 1080
QY 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 1081 ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGGAAGATGGGACTCGTCAT 1140
QY 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db 1141 GACCACGCCCATCTTACTGACTGCTCGGATATATGTTCTCGGAAGAAATGAGCCCTGTGAC 1200
QY 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrcArgSerCysPheThrIleAsn 420
Db 1201 ACTTTGGGATTTGCACCCCAAGTAGTGAATGTGTAGTAAATATCGCAGCTGCACGATTAT 1260
QY 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGlnSerGlyHisAsnPheGly 440
Db 1261 GAAGATACAGGTCTTGGACTGGCCTTCACCATTTGCCATGAGTCTGGACACAACTTTGGC 1320
QY 441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460

Db 1321 ATGATTTCATGATGGAGAAGGGAACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCT 1380
QY 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrcLeuHis 480
Db 1381 ACATTGGCAGGACGCAATGGAGTCTTCTCTGGTCAACCTCGAGCGCCAGTATCTACAC 1440
QY 481 LysPheLeuArgSerValLysMetProAlaLeu 491
Db 1441 AAATTTCTAAGATCAGTGAAATGCCAGCTCTC 1473

RESULT 3

US-09-930-872-3
; Sequence 3, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. 6448388el Human Proteases and Polynucleotides Encoding the San
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/09/930,872
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3675
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-930-872-3

Alignment Scores:
Pred. No.: 4,66e-234 Length: 3675
Score: 2592.00 Matches: 486
Percent Similarity: 99.39% Conservative: 3
Best Local Similarity: 98.78% Mismatches: 2
Query Match: 98.63% Indels: 1
DB: 3 Gaps: 0

US-10-804-457-2 (1-491) x US-09-930-872-3 (1-3675)

QY 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaLeuTrpMetLeuLeuAlaGln 20
Db 1 ATGAAGCCCGCGCGCGGATGGCGGGCTTGGCGGCGCTGTGGATGCTGTGGCGCAG 60
QY 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
Db 61 GTGGCCGAGCAGCACCCTGCGTCCCATGGACCCGCGAGCGGCGCTGGAGCCCG 120
QY 41 SerValProArgProProProAlaGluArgProGlyTrpMetGlyLysGlyGluTyrc 60
Db 121 AGCGTCCCGCTCTCTCCACCGCGAGCGCGGCTGGATGGAAAGGCGCAATAT 180
QY 61 AspLeuValSerAlaTyrcLysValAspHisArgGlyAspTyrcValSerHisGluIleMet 80
Db 181 GACCTGGTCTCTGCGCTACGAGGTTCACCAAGGCGCATTTAGTGTCCCATGAAATCATG 240
QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
Db 241 CACCATCAGCGCGGAGAGACAGTGGCCGTGTCCGAGGTGAGTCTCTTACCTTCGG 300
QY 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAla 120
Db 301 CTGAAAGGCTCCAGCAGCAGCTTCCACGTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360
QY 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 361 CCTGGCTTTATTTGTGCAGACGTTGGGAAAGACAGGCACCTAAGTCTGTGCAGACTTTACCG 420
QY 141 ProGluAspPheCysPheTyrcGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db 421 CCAGAGGACTTCTGTTCATCAAGGCTCTTTCGCGATCAACAGAAATCTCTCAGTGGCC 480

```
QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu 180
DB 481 CTTTCACACCTGCCAAGCGCTTGTGAGCATGATACGACAGAGAGCGAGATTACTTCCTA 540
QY 181 ArgProLeuProSerHisLeuSerThrPheLeuGlyArgAlaAlaGlnGlySerSerPro 200
DB 541 AGGCCACTCTCTTCCACACCTCTCATGGAACCTCGGACAGAGCTGCCCCAGGCGAGCTGCCA 600
QY 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
DB 601 TCCACAGTACTGACAGAGATCCACAGAGCCCATGCTCTCTGGGGCCAGTAGGCTCTG 660
QY 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
DB 661 GTGACCTCAAGGACATGGGAGCTGGCACATCAACCCCTGCACAGAGCGACTTTCGGCTG 720
QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
DB 721 GGACTGCCACAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCGCCCTCC 780
QY 261 LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
DB 781 AAGGAGACCTCTTTCATCTTCCAGATGAGTATAGCTTTCCTTACGGCATTAAGCGCTCT 840
QY 281 LeuLeuArgSerHisArgAsnGluLeuAsnValGluThrLeuValValAspLys 300
DB 841 CTTCTGAGGTCCATAGAAATGAAGAACTGAACGTGGAGACCTTGGTGGTGGTTCGACAA 900
QY 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrThrTyrValLeuThrIleLeuAsn 320
DB 901 AAGATGATGCAAAACCATGGCCATGAAATATACACACCTTACGTGCTCAGCATACTCAAC 960
QY 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyLysAsnIleAlaIleVal 340
DB 961 ATGATATCTGCTTATTCGAAGATGGAACAAATAGGAGGAAACATCAACATTGCATTGTA 1020
QY 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis 360
DB 1021 GGTCTGATTTCTTAGAAGATGAACAGCCAGCAGCTGGTGAATAGTCAACGACGACAC 1080
QY 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
DB 1081 ACCTTAAGTAGCTCTGCCAGTGGCAGTCTGGATGATGGGAAAGATGGGACTCGTCAT 1140
QY 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
DB 1141 GACCACGCCATCTTACTGACTGCTCGATATATGTTCTGGAAGAATGAGCCCTGTGAC 1200
QY 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
DB 1201 ACTTTGGGATTTGCCACCAATAGTGAATGTGTAGTAAATATCGCAGCTGCACGATTAA 1260
QY 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnGly 440
DB 1261 GAAGATACAGGTCTTGGACTGGCCCTTACCATTCGCCATGAGTCTGGACACAACTTTGGC 1320
QY 441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
DB 1321 ATGATTCATATGGAAGAGGAAACATGTGTAAAGTCCGAGGCAACATCATGTCCCTCT 1380
QY 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
DB 1381 ACATTGGCAGGACCAATGGAGTCTTCTCTGGTACCCCTGCACGCCCGCCAGTATCTACAC 1440
QY 481 LysPheLeuArgSerValLysMetProAlaLeu 491
DB 1441 AAATTTCTAAGCACCGCTCAAGCTATCTGCTCTG 1474
```

RESULT 4

US-10-217-774-3

; Sequence 3, Application US/1021774

; Patent No. 6734007

; GENERAL INFORMATION:

```
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. 6734007el Human Proteases and Polynucleotides Encoding the
; FILE OF INVENTION: Same
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/10/217,774
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US/09/930,872
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3675
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-217-774-3

Alignment Scores:
Pred. No.: 4,66e-234 Length: 3675
Score: 2592.00 Matches: 486
Percent Similarity: 99.39% Conservative: 3
Best Local Similarity: 98.78% Mismatches: 2
Query Match: 98.63% Indels: 1
DB: 4 Gaps: 0
```

US-10-804-457-2 (1-491) x US-10-217-774-3 (1-3675)

```
QY 1 MetLysProArgAlaArgGlyTyrArgGlyLeuAlaAlaLeuTrpMetLeuLeuAlaGln 20
DB 1 ATGAAGCCCGCGCGCGCGGATGGCGGGCTTGGCGCGCTGTGGATGCTGCTGGCGCAG 60
QY 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
DB 61 GTGCGCGAGCAGCACCTCGTGGCCATGGGACCCGCGAGCGCGAGCGCTGGGAGCCG 120
QY 41 SerValProArgProProProProAlaGluArgProGlyTyrMetGluLysGlyGluTyr 60
DB 121 AGCGTCCCGCGCTCTCTCCACCGCGAGCGCGCGCTGGATGGAAAGCGGCAATAT 180
QY 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
DB 181 GACCTGGTCTCTGCTTACAGGTTGACCACAGGGCGGATTACGTGTCATGAATCATG 240
QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
DB 241 CACCATCAGCGCGGAGAGAGCAGTGGCCGTGTCGAGGTTGAGTCTCTTCACCTTCGG 300
QY 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAla 120
DB 301 CTGAAAGGCTCCAGGCACGACTTCCAGTGATCTGAGGACTTCCAGCAGCTAGTGGCT 360
QY 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
DB 361 CCGTGGCTTTATTGTGACAGCGTTGGGAAAGACAGGCACTAAGTCTGTGACACTTTACCG 420
QY 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
DB 421 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTCGATTCACAGAACTCTCAGTGGCC 480
QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu 180
DB 481 CTTTCAACCTGCCAAGGCTTGTCCAGCATGATACGAAAGAGAGGAGCATTTACTTCTTA 540
QY 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
DB 541 AGGCCACTCTCTTCCACACCTCTCACTGGAACCTCGGACAGCTGCCCAAGCGAGCTGCCA 600
QY 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
DB 601 TCCACGTACTGTACAAGAGATCCACAGAGCCCCATGCTCTTGGGGCCGAGTGTCTGCTG 660
```

```
QY 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
DB 661 GTGACCTCAAGGACATGGGAGCTGGCACATCAACCCCTGCACAGCAGCAGCTTCGCTG 720
QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysIleValThrProGlnProPro 260
DB 721 GGACTGCCCAAAAGCAGCATTTCTGTGAAGACGCAAGAAATACATGCCAGCCTCC 780
QY 261 LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
DB 781 AAGGAAGACCTCTTCATCTTCCAGATGAGTATAAGTCTTGTTCAGGCATAGCGCTCT 840
QY 281 LeuLeuArgSerHisArgGlnGluLeuAenValGluThrLeuValValValAspLys 300
DB 841 CTTCTGAGTCCCATAGAAATGAAGAACTGAACGTGGAGACCTTGTGTGTGGTCGACAAA 900
QY 301 LysMetMetGlnAsnHisGlyHisGluAenIleThrTyrValLeuThrIleLeuAen 320
DB 901 AAGATGATGCAAAACCATGGCCATGAAATATCACCACCTAGCTGTCTCAGGATCTCAAC 960
QY 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyLysIleAsnIleAlaIleVal 340
DB 961 ATGGTATCTGCTTTATTCAAGATGCAACATAGGAGGAAACATCAACATTGCAATTGTA 1020
QY 341 GlyLeuIleLeuLeuAspGluGlnProGlyLeuValIleSerHisIleAspHis 360
DB 1021 GGTCTGATTCTTCTAGAGATGAACAGCCAGGACTGGTGAATAGTCAACAGCAGACCCAC 1080
QY 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
DB 1081 ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCAT 1140
QY 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
DB 1141 GACCAGCCATCTTACTGCTGTCTGGATATATGTCTCTGGAGAAATGAGCCCTGTGAC 1200
QY 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAen 420
DB 1201 ACTTTGGGATTTGCCACCATAGTGAATGTGTAGTAATATCGCAGCTGCACGATTAAT 1260
QY 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
DB 1261 GAAGATACAGGCTTGGACTGGCCCTTCCACATTTGCCCATGAGTCTGGACACAACTTTGGC 1320
QY 441 MetIleHisAspGlyGluGlyAenMetCysLysLysSerGluGlyAsnIleMetSerPro 460
DB 1321 ATGATTCATGATGGAGAGGAAACATGTGTAAAAGTCCGAGGGCAACATCATGCTCCCT 1380
QY 461 ThrLeuAlaGlyArgGlnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
DB 1381 ACATTGGCAGGAGCGCAATGGAGTCTTCTCTGCTCACCCTGCAGCGCCAGTATCTACAC 1440
QY 481 LysPheLeuArgSer-ValLysMetProAlaLeu 491
DB 1441 AAATTTCTAAGCACCCGCTCAAGCTATCTGCCTTG 1474

RESULT 5
US-09-930-872-5
; Sequence 5, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Fziddle, Carl Johan
; APPLICANT: Hilbun, Erln
; TITLE OF INVENTION: No. 6448388el Human Proteases and Polynucleotides Encoding the Sa
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/09/930, 872
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4042
```

```
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-930-872-5

Alignment Scores:
Pred. No.: 5,41e-234 Length: 4042
Score: 2592.00 Matches: 486
Percent Similarity: 99.39% Conservative: 3
Best Local Similarity: 98.78% Mismatches: 2
Query Match: 98.63% Indels: 1
DB: 3 Gaps: 0

US-10-804-457-2 (1-491) x US-09-930-872-5 (1-4042)
QY 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaAlaLeuTrpMetLeuLeuAlaGln 20
DB 99 ATGAAGCCCGCGCGCGCGGATGGCGGCTTGGCGGCGCTGTGGATGCTGTGGCGCAG 158
QY 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
DB 159 GTGGCGGAGCAGGCACCTGCGTGCATGGGACCCGCGCAGCGCAGCGCTGGGAGCCCG 218
QY 41 SerValProArgProProProAlaGluArgProGlyTrpMetGlyLysGlyGluTyr 60
DB 219 ACGTCCCCGCTCTCTCCACCGCGGAGCGCGCGCTGGATGGAAAGGGCGGAATAT 278
QY 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
DB 279 GACCTGGTCTCTGCTACGAGTTCACACACAGGGGCGATTACGTGTCCCATGAATCATG 338
QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
DB 339 CACCATCAGCGCGGAGAGAGCAGTGGCGGTTCGAGGTTGAGTCTCTTCACTTCGG 398
QY 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerLeuValAla 120
DB 399 CTGAAGGCTCCAGGACACGACTTCCACGTGGATCTGAGGACTTCCAGCAGCCTTAGTGGCT 458
QY 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
DB 459 CTGGCTTTATTTGTCAGACGTTGGGAAGACAGGCACTAAGTCTGTGCAGACTTTACCG 518
QY 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
DB 519 CCAGAGGACTTCTGTCTTATCAAGGCTCTTGGCGATCACACAGAAACTCTCTCAGTGGCC 578
QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu 180
DB 579 CTTTCAACCTGCCAAGGCTTGTCAAGCATGATACGAACAGAGAGGACGATTACTTCTCTA 638
QY 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyValArgAlaAlaGlnGlySerSerPro 200
DB 639 AGGCCACTTCTTCCACCTCTCATGGAAACTCGGACAGAGCTGCCCAAGCAGCTGCCCA 698
QY 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
DB 699 TCCACGTACTGTACAAGAGATCCACAGAGCCCATGCTCTCTGGGCGCCAGTGGAGTCTG 758
QY 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
DB 759 GTGACCTCAAGGACATGGGAGCTGGCACATCAACCCCTGCACAGCAGCAGCTTCCGCTG 818
QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysIleValThrProGlnProPro 260
DB 819 GGACTGCCCAAAAGCAGCATTTCTGTGAAGACGCAAGAAATACATGATGCCAGCCTCCC 878
QY 261 LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
DB 879 AAGGAAGACCTCTTCTATCTTGCAGATGAGTATAAGTCTTGTTCAGGCATAGCGCTCT 938
QY 281 LeuLeuArgSerHisArgGlnGluLeuAenValGluThrLeuValValValAspLys 300
DB 939 CTTCTGAGTCCCATAGAAATGAAGAACTGAACGTGGAGACCTTGTGTGGTGGTCGACAAA 998
```

Qy 301 LysMetMetGlnAenHisGlyHisGluAenIleThrThrTyrrValLeuThrIleLeuAen 320
 Db 999 AAGATGATGCAAAACCATGGCCATGAAATATACCCACCTACGTCTCACGATCTCAAC 1058
 Qy 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAenIleAenIleAlaIleVal 340
 Db 1059 ATGGTATCTGCTTTATTCAAAGATGGACAAATAGGAGGAACATCAATTCGCAATGTGA 1118
 Qy 341 GlyLeuIleLeuLeuGluAenGluInProGlyLeuValIleSerHisHisAlaAspHis 360
 Db 1119 GGTCTGATCTCTTCTAGAGATGAACAGCCAGGACTGGTGATAAGTCAACCGCAGACAC 1178
 Qy 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
 Db 1179 ACCTTAAGTAGCTCTGCGCAGTGGCAGTCTGGATTGATGGGGAAGATGGGACTCGTCAT 1238
 Qy 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAenGluProCysAsp 400
 Db 1239 GACCAAGCCATCTTACTGACTGGTCTGGATATATGTTCTTGGAGAGATGAGCCCTGTGAC 1298
 Qy 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrrArgSerCysThrIleAen 420
 Db 1299 ACTTTGGATTTGACCCATGAAGTGGATGTGTAGTAAATATCGCAGCTCGACGATTAAT 1358
 Qy 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAenPheGly 440
 Db 1359 GAAGATACAGTCTTGGACTGGCCTTCAACATTGCCATGAGTCTGGACACAACTTTGGC 1418
 Qy 441 MetIleHisAspGlyGluGlyAenMetCysLysLysSerGluGlyAenIleMetSerPro 460
 Db 1419 ATGATTCATGATGGAAGGAACATGTGTAAAGTCCGAGGCAACATCATGTCCCT 1478
 Qy 461 ThrLeuAlaGlyArgAenGlyValPheSerTrpSerProCysSerArgGlnTyrrLeuHis 480
 Db 1479 ACATTGGCAGGCAATGGAGTCTCTCTGTGTACCTTGACCGCCGCGATCTATACAC 1538
 Qy 481 LysPheLeuArgSer-VallysMetProAlaLeu 491
 Db 1539 AATTTCTAAGCACCGCTCAAGCTATCTGCCTTG 1572

RESULT 6

US-10-217-774-5
 ; Sequence 5, Application US/1021774
 ; Patent No. 6734007
 ; GENERAL INFORMATION:
 ; APPLICANT: Eridelle, Carl Johan
 ; APPLICANT: Hilbun, Erin
 ; TITLE OF INVENTION: Same
 ; TITLE OF INVENTION: Same
 ; FILE REFERENCE: LEX-0219-USA
 ; CURRENT APPLICATION NUMBER: US/10/217,774
 ; PRIOR FILING DATE: 2002-08-12
 ; PRIOR FILING DATE: US/09/930,872
 ; PRIOR FILING DATE: 2001-08-14
 ; PRIOR APPLICATION NUMBER: US 60/225,852
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 4042
 ; TYPE: DNA
 ; ORGANISM: homo sapiens
 US-10-217-774-5

Alignment Scores:
 Pred. No.: 5,41e-234 Length: 4042
 Score: 2592.00 Matches: 486
 Percent Similarity: 99.39% Conservative: 3
 Best Local Similarity: 98.78% Mismatches: 2
 Query Match: 98.63% Indels: 1
 DB: 4 Gaps: 0

US-10-804-457-2 (1-491) x US-10-217-774-5 (1-4042)
 Qy 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaAalaLeuTrpMetLeuLeuAlaGln 20
 Db 99 ATGAAGCCCGCGCGCGCGATGGCGGCTTGGCGGCGCTGTGGATGCTGTGGCGCAG 158
 Qy 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAalaAalaProGlySerPro 40
 Db 159 GTGGCCGAGCAGCACCCTCGTGGCGCATGGGACCCGAGCGCAGCGCTCGGAGCCCG 218
 Qy 41 SerValProArgProProProAlaGluArgProGlyTrpMetGluLysGlyGluTyrr 60
 Db 219 AGCGTCCCGCGTCTCTCCACCCCGGAGCGCGCGGCTGGATGAGAAAGCGCGAATAT 278
 Qy 61 AspLeuValSerAlaTyrrGluValAspHisArgGlyAspTyrrValSerHisGluIleMet 80
 Db 279 GACCTGGTCTCTCGCTACGAGTTGACCACAGGGCGGATACGTCTCCCATGAAATCATG 338
 Qy 81 HisHisGlnArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
 Db 339 CACCATCAGCGCGGAGAGAGAGTGGCGGTGTCGAGGTGAGTCTCTTACCTTCGG 398
 Qy 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAla 120
 Db 399 CTGAAAGGCTCCAGGCGCAGCTTCCACGTGGATCTGAGGACTTCCAGCAGCGCTAGTGCT 458
 Qy 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
 Db 459 CTTGGCTTTATGTGCAGAGCTGGGAAAGACAGGACATAAGTCTGTGCAGACTTTACCG 518
 Qy 141 ProGluAspPheCysPheTyrrGlnGlySerLeuArgSerHisArgAenSerSerValAla 160
 Db 519 CCAGAGGACTTCTGTCTTCTATCAAGGCTCTTTCGGATCACACAGAACTCTCAGTGGCC 578
 Qy 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrrPheLeu 180
 Db 579 CTTTCAACCTGCCAGGCTTGTTCAGGATGATACGAAACAGAGAGGAGGAGTACTTCTTA 638
 Qy 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyValArgAlaAalaGlnGlySerSerPro 200
 Db 639 AGGCCACTTCTTTCACACTCTCATGAAACTCGGAGAGCTGCCAGAGCTGCCAGAGCTGCCA 698
 Qy 201 SerHisValLeuTyrrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
 Db 699 TCCACGTACTGTACAGAGATCCACAGAGCCCCATGTCTCTGGGCGCAGTGGCTCTG 758
 Qy 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
 Db 759 GTGACCTCAAGGACATGGGAGCTGGCAGATCAACCCCTGCACAGCAGGAGCTTCGCTG 818
 Qy 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrrMetProGlnProPro 260
 Db 819 GGACTGCCCAAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCGCCAGCTCCC 878
 Qy 261 LysGluAspLeuPheIleLeuProAspGluTyrrLysSerCysLeuArgHisLysArgSer 280
 Db 879 AAGGAAGACCTCTTCTCATCTTCCAGATGAGTATAGTCTTCTTACGCGCATTAAGCGCTCT 938
 Qy 281 LeuLeuArgSerHisArgAenGluGluLeuAenValGluThrLeuValValAspLys 300
 Db 939 CTTCTGAGGTCCCATAGAAATAGAAAGCTGAACCTGGAGACCTTGGTGGTGGTGCACAAA 998
 Qy 301 LysMetMetGlnAenHisGlyHisGluAenIleThrThrTyrrValLeuThrIleLeuAen 320
 Db 999 AAGATGATGCAAAACCATGGCCATGAAATATACCCACCTACGTCTACGATCTCAAC 1058
 Qy 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyAenIleAenIleAlaIleVal 340
 Db 1059 ATGGTATCTGCTTTATTCAAAGATGGAAACATAGGAGGAACATCAACATTCGCAATGTGA 1118
 Qy 341 GlyLeuIleLeuLeuGluAenGluInProGlyLeuValIleSerHisHisAlaAspHis 360
 Db 1119 GGTCTGATCTCTTCTAGAGATGAACAGCCAGGACTGGTGTAGTAAAGTCAACCGCAGACAC 1178

361	ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis	380
1179	ACCTTAAGTAGCTTCTGCCAGTGCACTGTGGATTGATGGGGAAAGATGGGACTCGTCAT	1238
381	AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp	400
1239	GACCACGCCATCTTACTGACTGGTCTGGATATATGTTCTCGGAAGAATGAGCCCTGTGAC	1298
401	ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTrpArgSerCysThrIleAsn	420
1299	ACTTTGGGATTTGGACCCATAAGTGGAAATGTGTAGTAAATATGCGCAGCTGCACGATTAAAT	1358
421	GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly	440
1359	GAAGATACAGGCTCTTGGACTTGGCCCTTCACCAATTGCCCATGAGCTCGACACCACTTTGGC	1418
441	MetIleHisAspGlyGluGlyAsnMetCysAlaValysSerGluGlyAsnIleMetSerPro	460
1419	ATGATTTCATGATGAGAAGGGAACATGTGTAAAAAGTCCGAGGGCAACATCATGTGCCCT	1478
461	ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis	480
1479	ACATTGCGAGGACGCAATGGAGTCTTCTCTGTGTACCTGCGCCCGGACTATCTACAC	1538
481	LysPheLeuArgSerValLysMetProAlaLeu	491
1539	AAATTTCTTAAGCACCGCTCAAGCTATCTGCGCTTG	1572

RESULT 7

```

US-09-963-791-5
; Sequence 5, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Dotoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sdoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the Sa
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1407
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-963-791-5

```

Alignment Scores:	
Pred. No.:	3,956-58
Score:	720.50
Percent Similarity:	50.9%
Best Local Similarity:	35.8%
Query Match:	72.4%
DB:	4
Length:	1407
Matches:	168
Conservative:	71
Mismatches:	141
Indels:	89
Gaps:	15

US-10-804-457-2 (1-491) x US-09-963-791-5 (1-1407)

Qy	60	TyrAspLeuValSerAla	TyrGluValAspHis	ArgGlyAspTyrValSerHis	GluIle 79
Db	124	TACAGCTAACTAT	TCCAAATAAGGGTTGAT	CAAAATGGAGCATTTCTCA	CGCTTTCACGTTTACTGTG 183
Qy	80	MetHis-----	HisGlnArgArgAlaVal	-----AlaValSerGlu	93
Db	184	AAAAATGATAAACACT	CAAGGAGAAAGACGGAGTAT	TGGACCCCTATTGATTCACACAGCAGCA	243
Qy	94	ValGluSerLeuHis	LeuArgGluValSerArgHis	AspPheHisValAspLeuArg	113

Df 1177 ATTGGCTGGGTTTCAGCTTTTACCATGTGCATGAGATTGGTCACAATTTTTGGTATGAAC 1236

Qy 443 HisAspGlyGluGlyAsnMetCys-----LysLysSerGlu----GlyAsnIleMetSer 459
|||||
Df 1237 CATGATGNAATTGGAAATCTTTGTGGGACGAAAGGTCATGAAGCAGCAAACTTATGGCA 1296
|||||
Qy 460 ProThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeu 479
:::
Df 1297 GCTCACATTACTGCGAATAACCAATCCCTTTTTCCTGGTCTGCTTGCGATCGAGACTATC 1356
|||
Qy 480 HisLysPheLeuArgSerValLysMet 488
|||||
Df 1357 ACCAGCTTTCTAGAATTCTTTAAACTC 1383
|||

RESULT 8
US-09-963-791-9
; Sequence 9, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encod
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-963-791-9

Alignment Scores:
Pred. No.: 1,07e-57 Length: 1524
Score: 716.50 Matches: 168
Percent Similarity: 50.86% Conservative: 69
Best Local Similarity: 36.05% Mismatches: 140
Query Match: 27.26% Indels: 89
DB: 4 Gaps: 15

US-10-804-457-2 (1-491) x US-09-963-791-9 (1-1524)

Qy 60 TyrAspLeuValSerAlaTyrgluValAspHisArgGlyAspTyrValSerHisGluIle 79
|||||
Df 124 TACCAGCTAACCTATTCCAATAGGGTTGATCAAATGGAGCATTTCTCAGCTTTACTGTG 183
|||||
Qy 80 MetHis-----HisGlnArgArgArgAlaVal-----AlaValSerGlu 93
:::
Df 184 AAAAAATGATAACACTCAAGGAGAGACGGAGTAGTGACCTTATGTATCCACAGCAGCA 243
|||||
Qy 94 ValGluSerLeuHisLeuArgGlnLysGlySerArgHisAspPheHisValAspLeuArg 113
:::
Df 244 GTATCTAGTTATTTTTTAATCTTTTCAGCCTATGGCAAGCACATTTTCATCTAAACTTGACT 303
|||||
Qy 114 ThrSerSerSerLeuValAlaProGlyPheIleValGlnThrLeuGlyLysThrGlyThr 133
:::::
Df 304 CTCACACACAGATTTTGTGTCCAAACATTTTACATAGATATATGGGGGAAAGATGGA--- 360
|||||
Qy 134 LysSerValGlnThrLeuProProGlu-----AspPheCysPheTyr 147
|||||
Df 361 -----CCCAGTGGGAAACATGATTTTTTTTACAGCAACTGTCATTAC 399
|||||
Qy 148 GlnGlySerLeuArgSerHisAtqAsnSerSer---ValalaleuSerThrCysGlnGly 166
|||||
Df 400 ACAGATATTTGCCAAGTCAACGTAGTACATTAAGTGCTTTAAGCAACTGTGTGGG 459

US-09-963-791-11
; Sequence 11, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedreich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the Sa
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1770
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-963-791-11

Alignment Scores:
Pred. No.: 1.35e-57 Length: 1770
Score: 716.50 Matches: 168
Percent Similarity: 50.86% Conservative: 69
Best Local Similarity: 36.05% Mismatches: 140
Query Match: 27.26% Indels: 89
DB: 4 Gaps: 15

US-10-804-457-2 (1-491) x US-09-963-791-11 (1-1770)

QY 60 TyrAspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluLeu 79
DB 124 TACCAGCTAACTATTCCCAATAAGGGTGTATCAAAATGGAGCAATTTCTCAGCTTTACTGTG 183
QY 80 MetHis-----HisGlnArgArgArgAlaVal-----AlaValSerGlu 93
DB 184 AAAATGATAACACTCAAGGAGAGACGGAGTATGGACCCCTATTGATCCACAGCAGCA 243
QY 94 ValGluSerLeuHisLeuArgLeuGlySerArgHisAspPheHisValAspLeuArg 113
DB 244 GTATCTAAGTTATTTTAAACTTTTCAGCCTATGCGACCACTTTTCATCTAAACTTGACT 303
QY 114 ThrSerSerLeuValAlaProGlyPheLeuValGlnThrLeuGlyLysThrGlyThr 133
DB 304 CTCACACAGATTTTGTGTCCAAACATTTTACAGTAGAATATTGGGGGAAAGATGGA--- 360
QY 134 LysSerValGlnThrLeuProGlu-----AspPheCysPheTyr 147
DB 361 -----CCCAGTGGAAACATGATTTTATAGACAACTGTCTATTAC 399
QY 148 GlnGlySerLeuArgSerHisArgHisSer-----ValAlaLeuSerThrCysGlnGly 166
DB 400 ACAGGATATTTGCAAGATCAAGTAGTACAACTAAAGTGGCTTAAAGCAACTGTGTGGG 459
QY 167 LeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeuArgProLeuProSer--- 185
DB 460 TTGCATGGTGTATTGCTACAGAAAGATGATTTTATCGAACCTTTTAAAGATAC 519
QY 186 -----HisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerPro 200
DB 520 ACAGAGGATTCAGCATTTTATGTAATAATGGCCAC----- 558
QY 201 SerHisValLeuTyrLysArgSerThr-----GluProHis 212
DB 559 CCTCATGTTATTACAAAAGCTGCCCTTCAACACGACATCTGTATGATCACTCTCAT 618
QY 213 AlaProGlyAlaSerGluValLeuValThrSerArgThrTrpGluLeuAlaHisGlnPro 232
DB 619 TGT---GGGGTTTCGATTTCACAAAGAGTGGCAACCTTTGGTGGCTGAATGACATCC 675

QY 233 LeuHisSerSerAspLeuArgLeuGlyLeuProGlnLysGlnHisPheCysGlyArgArg 252
DB 676 ACTGTTCTTCTATTCTACTACCAATTAAACACACATATCCACCAC----- 720
QY 253 LysLysTyrMetProGlnProLysGluAspLeuPheLeuProAspGluTyrLys 272
DB 720 ----- 720
QY 273 SerCysLeuArgHisLysArgSerLeuLeuArgSerHisArgAsnGluGluLeuAsnVal 292
DB 721 -----AGACAGAGAGATCAGTG-----AGCATTGAACGGTTTGTG 756
QY 293 GluThrLeuValValAspLysLysMetMetGlnAsnHisGlyHisGlnAsnLeuThr 312
DB 757 GAGACATTGGTAGTGGCAGACAAATGATGTGGTGGCTACCATGGCGCGCAAGACATTGAA 816
QY 313 ThrTyrValLeuThrIleLeuAsnMetValSerAlaLeuPheLysAspGlyThrIleGly 332
DB 817 CATTACATTTTGGTGTGATGAATATTGTTGCCAACTTTTACCGTGATTTCCAGCCTAGGA 876
QY 333 GlyAsnIleAsnIleAlaIleValGlyLeuLeuLeuAspGluGlnProGlyLeu 352
DB 877 AACGTTGTGAATATTATAGTGGCCCGCTTAATTGTTCTACAGAAAGATCAGCAAACTTG 936
QY 353 ValIleSerHisHisAlaAspHisThrLeuSerSerPheCysGlnTrpGlnSerGlyLeu 372
DB 937 GAGATAAACCACTGATGACAGCAAGTCCCTCGATAGCTTCTGTAAATGGCAGAAATCCATT 996
QY 373 MetGly-----LysAspGly-----ThrArgHisAspHis 382
DB 997 CTCTCCCAACCAAGTGTGTAACCACTTCCAGAAAAATGGGATTTGCCACCACCATAT 1056
QY 383 AlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAspThrLeu 402
DB 1057 GCAGTCTTATTATGATATGATATCTGCATTAATAAATTAACCCCTGTGGAACTG 1116
QY 403 GlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsnGluAsp 422
DB 1117 GCCTTGGCCTCTGTGGCTGGAATGTGTGAGCCTGAAGAGAGCTGCAGCATTAATGAAGAC 1176
QY 423 ThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGlyMetIle 442
DB 1177 ATTGGCCTGGGTTCAGCTTTTACCATTGCATGATGATGGTGGTTCACAAATTTTGGTATGAAC 1236
QY 443 HisAspGlyGluGlyAsnMetCys-----LysLysSerGlu---GlyAsnIleMetSer 459
DB 1237 CATGATGGAATGGAAATTTCTTGGGGAGAAAGCTCATGAAGCAGCAAACTTATGGCA 1296
QY 460 ProThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeu 479
DB 1297 GCTCACATTACTGCGAATACCAATCCTTTTCTGTGCTGTGCTGCGAGCTACATC 1356
QY 480 HisLysPheLeuArgSer 485
DB 1357 ACCAGCTTTCTAGATTCA 1374

RESULT 10
US-09-963-791-1
; Sequence 11, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedreich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the San
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769

; PRIOR FILING DATE: 1999-12-09
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 2727
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-963-791-1

Alignment Scores:
 Pred. No.: 2,65e-57 Length: 2727
 Score: 716.50 Matches: 168
 Percent Similarity: 50.86% Conservative: 69
 Best Local Similarity: 36.05% Mismatches: 140
 Query Match: 27.26% Indels: 89
 DB: 4 Gaps: 15

US-10-804-457-2 (1-491) x US-09-963-791-1 (1-2727)

```

QY 60 TyrAspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIle 79
DB 124 TACCAGCTAACTATTCCAAATAGGGTTGATCAAATGGAGCATTTCTAGCTTTACTGCG 183
QY 80 MetHis-----HisGlnArgArgArgAlaVal-----AlaValSerGlu 93
DB 184 AAAAAATGATAAACACTCAAGGAGGAGACGAGTATGGACCTATTGATCCACAGCAGGCA 243
QY 94 ValGluSerLeuHisLeuLeuLysGlySerArgHisAspPheHisValAspLeuArg 113
DB 244 GTATCTAAGTATTTTAACTTTCAGCCTTGGCAGCATTTCATCTAAACTTGACT 303
QY 114 ThrSerSerSerLeuValAlaProGlyPheIleValGlnThrLeuGlyLysThrGlyThr 133
DB 304 CTCACACAGATTTTGTGTCCAAACATTTTACAGTAGAATATTGGGGGAAAGATCGA--- 360
QY 134 LysSerValGlnThrLeuProGlu-----AspPheCysPheTyr 147
DB 361 -----CCCAAGTGGAACATGATTTTTTAGACAACCTGTCAATTAC 399
QY 148 GlnGlySerLeuArgSerHisArgAsnSerSer---ValAlaLeuSerThrCysGlnGly 166
DB 400 ACAGGATATTTCGAAGATCAACGTAGTACAATAAGTGGCTTTAAGCAACTGTGTGGG 459
QY 167 LeuSerGlyMetIleArgThrCluGluAlaAspTyrPheLeuArgProLeuProSer--- 185
DB 460 TTGCATGGTGTTATTCACAGAGATGAAGAGTATTTTATCGAACCTTTAAAGAATACC 519
QY 186 -----HisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
DB 520 ACAGAGGATTCACAGCATTTTAGTTATGAAATGGCCAC----- 558
QY 201 SerHisValLeuTyrLysArgSerThr-----GluProHis 212
DB 559 COTCATGTTATTTTACAAAAGTCTGCCCTTCAACAACGACATCTGTATGATCACTCTCAT 618
QY 213 AlaProGlyAlaSerGluValLeuValThrSerArgThrTrpGluLeuAlaHisGlnPro 232
DB 619 TGT---GGGGTTTCGATTTCCAGAAAGTGGCAACCTTGGTGGCTGAATGACACATCC 675
QY 233 LeuHisSerSerAspLeuArgLeuGlyLeuProGlnLysGlnHisPheCysGlyArgArg 252
DB 676 ACTGTTTCTTATTCATCACTAACCAATTAACAACACATATCCACAC----- 720
QY 253 LysLysTyrMetProGlnProProLysGluAspLeuPheLeuProAspGluTyrLys 272
DB 720 ----- 720
QY 273 SerCysLeuArgHisLysArgSerLeuLeuArgSerHisArgAsnGluGluLeuAsnVal 292
DB 721 -----AGACAGAGAGATCAGTG-----ACCATTAACCGTTTGTG 756
QY 293 GluThrLeuValValAspLysLysMetMetGlnAsnHisGlyHisGluAsnIleThr 312
DB 756 -----
  
```

```

DB 757 GAGACATTGGTAGTGGCAGCAAAATGATGGTGGCTACCATCGCGCAAGACATTGAA 816
QY 313 ThrTyrValLeuThrIleLeuAsnMetValSerAlaLeuPheLysAspGlyThrIleGly 332
DB 817 CATTACATTTTGAGTGTGATGATTAITTTGTTCCAAACITTTACCGTATTCCAGCCTAGGA 876
QY 333 GlyAsnIleAsnIleAlaIleValGlyLeuIleLeuLeuGluAspGlnProGlyLeu 352
DB 877 AACGTTGTGAATATTATAGTGGCCGCTTAATTGTTCTCACAGAAAGATCAGCAAACTTG 936
QY 353 ValIleSerHisHisAlaAspHisThrLeuSerSerPheCysGlnTrpGlnSerGlyLeu 372
DB 937 GAGATAAACCCATCGCAGCAAGTCCCTCGATAGCTTCTGTAAATGGCAGAAATCCATT 996
QY 373 MetGly-----LysAspGly-----ThrArgHisAspHis 382
DB 997 CTCCTCCCAACCAAGTATGATGAAACACCATTCACGAAATGGATGGTCCACCATGATTA 1056
QY 383 AlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAspThrLeu 402
DB 1057 GCAGTTCTTATTACTAGATATGATATCTGCACATTATAAAAATAAGCCCTGTGGAACACTG 1116
QY 403 GlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsnGluAsp 422
DB 1117 GGCTTGGCCTCTGTGGCTGGATGTGTGAGCCTGAAAGGAGCTGCAGCATTAATGAAGAC 1176
QY 423 ThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGlyMetIle 442
DB 1177 ATTGGCCTGGTTCAGCTTTTACCATTCACATGAGATGGTTCACATTTTGGTATGAAAC 1236
QY 443 HisAspGlyGluGlyAsnMetCys-----LysLysSerGlu---GlyAsnIleMetSer 459
DB 1237 CATGATGAATTTGAAATTTCTTGTGGACGAAAGGCTCATGAAGCAGCAAACTTATGGCA 1296
QY 460 ProThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeu 479
DB 1297 GCTCACATTACTCGAATACCAATCTTTTCTGCTGGTCTGCTGCTGAGTGCAGTACATC 1356
QY 480 HisLysPheLeuArgSer 485
DB 1357 ACCAGCTTTCTAGATTCA 1374

```

RESULT 11

US-09-369-364A-6
 ; Sequence 6, Application US/09369364A
 ; Patent No. 6391610
 ; GENERAL INFORMATION:
 ; APPLICANT: Apte, Suneel
 ; APPLICANT: Hurskainen, Tiina L.
 ; APPLICANT: Hirohata, Satoshi
 ; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
 ; FILE REFERENCE: 26473/4007/10-30-00
 ; CURRENT APPLICATION NUMBER: US/09/369,364A
 ; CURRENT FILING DATE: 1999-08-06
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 3218
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens ADAMTS-7
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (13) .. (3003)
 ; US-09-369-364A-6

Alignment Scores:
 Pred. No.: 4.76e-57 Length: 3218
 Score: 715.00 Matches: 176
 Percent Similarity: 49.69% Conservative: 63
 Best Local Similarity: 36.59% Mismatches: 176
 Query Match: 27.21% Indels: 66
 DB: 3 Gaps: 12

US-10-804-457-2 (1-491) x US-09-369-364A-6 (1-3218)

Qy	32	ProLaAlaAlaAlaProGlySerProSerValProArgProProProProAlaGluArg	51
Db	1	CGGGTTCTGCGATCGCCGGCGGCCCAAGT---CCCGCAGCCCGCGCGCTTTGCTGGCG	57
Qy	52	-----ProGlyTrp	54
Db	58	CCCTCTCTCTCTCTCTGCGCTCTGCGTCCCGCGGCCCGCGACCGCCACACGAGCGT	117
Qy	55	MetGluLysGlyGlu-----TyrAspLeuValSerAlaTyrGluValAspHisArgGly	72
Db	118	GCAACCGAGGCGCGCGCGACTGCATCTGTCACCGCTCGAGTTCGACGCGGGGGC	177
Qy	73	AspTyrValSerHisGluIleMethHisGlnArgArgArgAlaValAlaValSer	92
Db	178	TCCTTCTGTCTCTACGAGCTGTGGCCCGCGCACTGCGCAGCGGCGATGTAATCTGTGCGC	237
Qy	93	GluValGluSerLeuHisLeuArgLeuLysGlySerArgHisAspPheHisValAspLeu	112
Db	238	CGAGACGCGCCCGCTTCTACGAGCTACAAATACCGCGCGCGAGCTCGCTTCAACCTG	297
Qy	113	ArgThrSerSerLeuValAlaProGlyPheIleValGlnThrLeuGlyLysThrGly	132
Db	298	ACCGCAATCAGCACTCTGTGGCCCGCGCTTTGAGGAGACGCGCGCGCGCGCGCGC	357
Qy	133	ThrLysSerValGlnThrLeuProGluAspPheCysPheTyrGlnGlySerLeuArg	152
Db	358	CTGGCGCGCGCACATCCGGGCCCACACCCCGGCTGCCACCTGCTTGGCGAGGTGCAG	417
Qy	153	SerHis-----ArgAsnSerSerValAlaLeuSerThrCysGlnGlyLeuSerGlyMet	170
Db	418	GACCTCGAGCTCGAGGTGGCTGTGGCGGCATCAGCGCTTCGACGGCGCTGAAGAGTGTG	477
Qy	171	IleArgThrGluGluAlaAspTyrPheLeuArgProLeuProSerHisLeuSerTrpLys	190
Db	478	TTCCAGCTCTCAACGAGGACTACTTCATGAGCCCTGGACAGT-----	522
Qy	191	LeuGlyArgAlaAlaGlnGlySerSerProSerHisValLeuTyrLysArgSerThrGlu	210
Db	523	---GCCCGCGCGCGCTGGCCACGCCCGCCCATGTGGTGTACAAGCGT-----	570
Qy	211	ProHisAlaProGlyAlaSerGluValLeuValThrSerArgThrTrpGluLeuAlaHis	230
Db	571	---CAGGCGCGCGAG-----AGGCTGGCGACAG	594
Qy	231	GlnProLeuHisSerSerAspLeuArgLeuGlyLeuProGlnLysGlnHisPheCysGly	250
Db	595	CGGGTGATTCAGGTGCTCCAAAGCACCTGTGGAGTGCAGGTGTAACCAAGCTGGAGTCT	654
Qy	251	ArgArgLysLysTyrMetProGlnProProLysGluAspLeuPheIleLeuProAspGlu	270
Db	655	CGACGGGAGCGTGG-----GAG	672
Qy	271	TyrLysSerCysLeuArgHisLysArgSerLeuLeuArgSerHisArg-----	287
Db	673	CAGCGCGCAGTGTGGCGCGGCCACGG---CTGAGCGCTCTACACGCGGCTGGCTCAGC	729
Qy	288	GluGluLeuAsnValGluThrLeuValValAspLysLysMetMetClnAsnHisGly	307
Db	730	AAAGAGAGATGGTGTGTGAGACCCCTGGTAGCTGATGCCAAATCGTGATACCACCGA	789
Qy	308	HisGluAsnIleThrThrTyrValLeuThrIleLeuAsnMetValSerAlaLeuPheLys	327
Db	790	CAGCGCGAGGTGTGAGAGCTATGTGTGACCATCATGAACATGGTGGCTGGCTGTTTCA	849
Qy	328	AspGlyThrIleGlyGlyAsnIleAlaIleValGlyLeuIleLeuLeuGluAsp	347
Db	850	GACCCAGCATGGGACCCCATCCATCACCATTGTGGCGCTGGTCTGTCTGCTGGAAGT	909
Qy	348	GluGlnProGlyLeuValIleSerHisAlaAspHisThrLeuSerSerPheCysGln	367
Db	910	GAGGAGGAGGACCTAAAGATCAGGACCATGAGCAGCAACACCTGAAGAGCTTCTGCAAG	969

Qy		368	TripGlnSerClyLeuMetGlyLysAspGlyThr-----ArgHisaspHisaIalle	384
Db		970	TGGCAGAAAGCATCAACATGAAGGGGGGTGCCCATCCCTCGCACCATGCATCGGCATC	1029
Qy		385	LeuLeuthrGlyLeuAaspileCyserSertTpLysAsnGluProCySaspThrrLeuGlyPhe	404
Db		1030	CTGCTACCAGAAGGACCTGTGTGAGCCATGAA CGGCCCCTGTGAGACCCCTTGGGACTG	1089
Qy		405	AlaProileSerGlyMetCySserLystrYzrArgSerCysthrlleasnGluAspThrrGly	424
Db		1090	TCCCATGTGGCGGCATGTSCCAGCCGCACTGCAGCATCAACAGGACACCGGCG	1149
Qy		425	LeuClyLeualaphethrllealAHisGluSerGlyHiShanPhedgylMetlleHisAsp	444
Db		1150	CTGCGGTGGCCCTTCATCTTAGCCCCAGACTCGGGCACAGATTTTGGCATTCACAGATGAC	1209
Qy		445	GlyGluclyAsnMetCvstLysLys-----SerGluGlyAsnilleMetSerProThrrLeu	462
Db		1210	GGAAGCGGCANTGACTGTGAGCCGGTGGGAAGACGACTTTTCATGTCTCCACAGCTC	1269
Qy		463	AlaGlyArgAsnGlyValPheSertTrpSerProcysSerargGIntyrLeuHisLysPhe	482
Db		1270	CTGTAGACACCGCTCCCTTCACCTTGTCCTCCGTGCAGCGCCAGTATATCACACAGGTTTC	1329
Qy		483	Leu 483	
Db		1330	Ctt 1332	

RESULT 12

```

US-09-369-364A-4
; Sequence 4, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2848
; TYPE: DNA
; ORGANISM: Homo sapiens ADAMTS-6
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (22)..(2601)
; NAME/KEY: misc feature
; LOCATION: (1369)
; OTHER INFORMATION: n = C
; NAME/KEY: misc feature
; LOCATION: (1620)
; OTHER INFORMATION: n=C
US-09-369-364A-4

```

Alignment Scores:		
Pred. No.:	2,786-56	Length: 2848
Score:	706.00	Matches: 166
Percent Similarity:	50.8%	Conservative: 70
Best Local Similarity:	35.7%	Mismatches: 139
Query Match:	26.8%	Indels: 89
DB:	3	Gaps: 14

US-10-804-457-2 (1-491) x US-09-369-364A-4 (1-2848)

Oy TyrAspLeuValSerAlaTyrGluValAspHisArgGlyAspTy rValSerHisGluIle 79
 ||||| ||||| ||||| ||||| ||||| ||||| ::::: ||||| ::
Dd TACCAGCTAACATTCCCAATAAGGGTTGCATCAAATGGAGCATTTTCACGTTTACTGTG 204

Ov Methis-----HisGLPArqArqArqAval-----AlavalSerGlu 93

205	Db	205	AAAAAGTATAAACCTCAAGGAGAGAGCGGAGTATTGACCTATTGATCCACAGCAGGCA	264
94	Qy	94	ValGluSerLeuHisLeuArgLeuLeuGlySerArgHisAspPheHisValAspLeuArg	113
265	Db	265	GTATCTAAGTTATTTTAAACCTTCAGCCTATGGCAAGCACATTTTCATCTAAACTTGACT	324
114	Qy	114	ThrSerSerSerLeuValAlaProGlyPheIleValGlnThrLeuGlnGlyThrGlyThr	133
325	Db	325	CTCAACACACAGATTTGTGTCCAAACATTTTACAGTAGAATATTTGGGGGAAGATGA	381
134	Qy	134	LysSerValGlnThrLeuProProGlu-----AspPheCysPheTyr	147
382	Db	382	-----CCCCAGTGGAAACATGATTTTTTTAGACAACTGTCTATTAC	420
148	Qy	148	GlnGlySerLeuArgSerHisArgAenSerSer---ValAlaLeuSerThrCysGlnGly	166
421	Db	421	ACAGGATATTTGCAAGATCAACGTAGTACAACTAAAGTGGCTTTAAGCAACTGTGTGGG	480
167	Qy	167	LeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeuArgProLeuProSer---	185
481	Db	481	TTGCATGGTGTATTGCTACAGAAGATGAAGAGTATTTTATCGAACCTTTAAGAATACC	540
186	Qy	186	-----HisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro	200
541	Db	541	ACAGAGGATTCCAAGCATTTTAGTTATGAAATGGCCAC-----	579
201	Qy	201	SerHisValLeuTyrIysArgSerThr-----GluProHis	212
580	Db	580	CCTCATGTTATTTCAAAAGAGTCTCCCTTCAACACGACATCTGTATGATCACTCTCAT	639
213	Qy	213	AlaProGlyAlaSerGluValLeuValThrSerArgThrTrpGluLeuAlaHisGlnPro	232
640	Db	640	TGT---GGGGTTTCGGATTTCACAAGAGTGGCAACCTTGTGGCTGGAATGACACTCCA	696
233	Qy	233	LeuHisSerSerAspLeuArgLeuGlyLeuProGlnLysGlnHisPheCysGlyArgArg	252
697	Db	697	CTGTTTCTTATTCACTACCAAAATAACAACACATATCCACCAC-----	741
253	Qy	253	LysLysTyrMetProGlnProProLysGluAspLeuPheIleLeuProAspGluTyrLys	272
741	Db	741	-----	741
273	Qy	273	SerCysLeuArgHisLysArgSerLeuLeuArgSerHisArgAenGluLeuAsnVal	292
742	Db	742	-----AGACAGAAGAGATCAGT-----AGCATTGAAACGGTTGTG	777
293	Qy	293	GluThrLeuValValAspLysLysMetMetGlnAsnHisGlyHisGluAsnIleThr	312
778	Db	778	GAGACATTTGTTGGCAGACAAATGATGGTGGCTACCATGGCCGCAAGACATTGAA	837
313	Qy	313	ThrTyrValLeuThrIleLeuAsnMetValSerAlaLeuPheLysAspGlyThrIleGly	332
838	Db	838	CATTACATTTTGTAGTGTGATGATATTGTTGCCAACTTTACGTGATTCAGCCCTAGGA	897
333	Qy	333	GlyAsnIleAsnIleAlaIleValGlyLeuIleLeuLeuGluAspGluGlnProGlyLeu	352
898	Db	898	AACGTTGTGAATATTATGATGGCCCGCGCTAATTTGTTCTCACAGAGATCAGCAAACTG	957
353	Qy	353	ValIleSerHisAlaAspHisThrLeuSerSerPheCysGlnTrpGlnSerGlyLeu	372
958	Db	958	GAGATAAACACCATTGCAGACAAATCCCTCGTAGCTTCTGTAATGGCAGAAATCCATT	1017
373	Qy	373	MetGly-----LysAspGly-----ThrArgHisAspHis	382
1018	Db	1018	CTCTCCCAAGTATGGAACACCATTCAGAAAAATGGGATTTGCCACCACCATTAAT	1077
383	Qy	383	AlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAspThrLeu	402
1078	Db	1078	GCAGTTCTTATTACTAGATATGATATCTGCATTTATAAAATAACCCCTGTGGAACTG	1137
403	Qy	403	GlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsnGluAsp	422

```
QY 108 PheHisValAspLeuArgThrSerSerSer-IeuValAlaProGlyPheIleValGlnTh 127
Db 5050 CTGCGCTTCAACCTGACCCCAATCAAGCACCTGCTGGCGCCGGCTTTGTGAGCGAGAC 4991
QY 127 rLeuGlyLysThrGlyThrLysSerValGlnThrLeuProGluAspPheCysPheTy 147
Db 4990 GCGCGCGCGCGCGCGCTGGCGCGCGCACATCCGGCGCCACACCCCGCGCTGCCACCT 4931
QY 147 rGlnGlySerLeuArgSerHis-----ArgAenSerSerValAlaLeuSerThrCysG 165
Db 4930 GCTTGGCGAGTGCAGGACCCCTGAGCTCGAGGGTGGCTGGCGCCATCAGCGCTGCCA 4871
QY 165 nGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeuArgProLeuProSe 185
Db 4870 CGGCTTGAAGAGTGTCTTCAGCTCTCCAAACGAGGACTACTTCATTGAGCGCCCTGGACAG 4811
QY 185 rHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerProSerHisValLeuTy 205
Db 4810 T-----GCCCCGCGCGCGCTGGCGCCAGCCAGCCCGCTGGGTGTGTA 4769
QY 205 rLysArg-----SerThrGluProHisAlaPr 214
Db 4768 CAAGCTCAGGCCCCGAGGAGGTGCACAGCGGGGTGATTCAGTGCTCCAGACACCTG 4709
QY 214 oGlyAlaSer-----GluValLeuValThrSerArgThrTrpGluLeuAlaHisG 231
Db 4708 TGGAGTGAAGTGTATCCAGAGCTGAGTCTCGACGGGAGCGTGGGAG-----4660
QY 231 nProLeuHisSerSerAspLeuArgLeuGlyLeuProGlnLysGlnHisPhe-CysGlyA 251
Db 4659 -----CAGCGCAGCAGTGGCGCGGC 4637
QY 251 rArgLysLysTyMetProGlnProProLysGluAspLeuPheIleLeuProAspGluT 271
Db 4636 CACGGCTGAGCGTCTA-----4620
QY 271 yLysSerCysLeuArgHisLysArgSerLeuLeuArgSerHisArgAenGluGluLeuA 291
Db 4619 -----CACCAGCGTGGTCT-----AGCAAGAGAGAAGT 4592
QY 291 snValGlnThrLeuValValAspLysMetMetGlnAenHisGlyHisGluAenI 311
Db 4591 GGGTGGAGACCTGGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 4532
QY 311 leThrThrTrpValLeuThrIleLeuAsnMetValSerAlaLeuPheLysAspGlyThrI 331
Db 4531 TTGAGAGCTATGTGCTGACCATCATGACATGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 4472
QY 331 leGlyGlyAsnIleAenIleAlaIleValIcLylLeuIleLeuLeuGluAspGluInProG 351
Db 4471 TTGGGAACCCATCCACATCACCATTGTGGCTGTGCTGTGCTGTGGAAGATGAGGAGG 4412
QY 351 lYeuValIleSerHisHisAlaAspHisThrLeuSerSerPheCysGlnTrpGlnSerG 371
Db 4411 ACCTAAGATCATCGCACCATTCAGACACACCCCTGAAGAGCTTCTCAAGTGGCAGAAA 4352
QY 371 lYeuMetGlyLysAspGlyThr-----ArgHisAspHisAlaIleLeuLeuThrG 388
Db 4351 GCATCAACATGAAGGGGATGCCATCCCTGTGCACCATCATGACACTGCCATCTGCTCACC 4292
QY 388 lYeuAspIleCysSerTrpLysAenGluProCysAspThrLeuGlyPheAlaProIles 408
Db 4291 GAAAGACCTGTGTGACGACCATGAACCGCCCTGTGAGACCTTGGGACTGTCCCATGTGG 4232
QY 408 eGlyMetCysSerLysTyArgSerCysThrIleAenGluAspThrGlyLeuGlyLeuA 428
Db 4231 CGGGCATGTGCCAGCGCCAGCTGCGATCAACGAGGACACCGGCGCTGCCCTGG 4172
QY 428 lApheThrIleAlaHisGluSerGlyHisAsnPheGlyMetIleHisAspGlyGluGlyA 448
Db 4171 CCTTACTGTAGCCCAACGAGCTCGGGCAGCATGTTTGGCATTCAGATGACGGAAGCGGCA 4112
```

```
QY 448 snMetCysLysLys-----SerGluGlyAsnIleMetSerProThrLeuAlaGlyArgA 466
Db 4111 ATAGCTGTGAGCCCGTGGGAAACGACACCTTTTCATCATGTCTCCACAGCTCTGTACGACG 4052
QY 466 snGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHisLysPheLeu 483
Db 4051 CGCTCTCCCTCACCTGGTCCCGCTGACGCGGCGCATATATCACCAGGTTCTT 3999

RESULT 14
US-09-981-953A-1
; Sequence 1, Application US/09981953A
; Patent No. 6689599
; GENERAL INFORMATION:
; APPLICANT: RACIE, LISA A.
; APPLICANT: TWINE, NATALIE C.
; APPLICANT: AGOSTINO, MICHAEL J.
; APPLICANT: WOLFMAN, NEIL
; APPLICANT: MORRIS, ELISABETH A.
; TITLE OF INVENTION: NOVEL AGGREGANASE MOLECULES
; FILE REFERENCE: 08702.0075-00000
; CURRENT APPLICATION NUMBER: US/09/981,953A
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/242,317
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3766
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Nucleotide
; OTHER INFORMATION: sequence of the aggreganase molecule
US-09-981-953A-1

Alignment Scores:
Pred. No.: 4,22e-55 Length: 3766
Score: 695.50 Matches: 180
Percent Similarity: 50.40% Conservative: 72
Best Local Similarity: 36.00% Mismatches: 190
Query Match: 26.46% Indels: 59
DB: 4 Gaps: 14

US-10-804-457-2 (1-491) x US-09-981-953A-1 (1-3766)
QY 5 AlaArgGlyTyrArgGlyLeuAlaAlaLeuTyrMetLeuLeuAlaGlnValAlaGluGln 24
Db 216 ACCCGGGCTGGGAGAGACATGT-----GGACACGTGGGCTCTATG 257
QY 25 AlaProAlaCysAlaMetGlyProAlaAlaAlaAlaProGlySer---ProSerVal--- 42
Db 258 GCTCCCGCTGCCAGATCCT-CCGCTGGGCGCTCGCCCTGGGCTGGGCTCATGTTTGA 316
QY 43 ProArgProProProProAlaGluArgProGlyTyrMetGluLysGlyGlu-TyrAspLe 62
Db 317 GGTCAACGACGCGCTCCCGTCTCAAGATGAGTTCCTGCTCCAGTCTGAGAGCTATGAGAT 376
QY 62 uValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMethHisH 82
Db 377 CGCTTCCCGCCCGCGGTGGACCAACCGGGGACGTGCTGGGCTTCTGCCACCTCTCTCC 436
QY 82 sGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArgLeuTy 102
Db 437 CCGGAGGACGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 496
QY 102 sGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAlaProGl 122
Db 497 CTCGCCACGACCCCACTTCTGCTGAACCTGACCCGCGAGCTCCCGTCTACTGCGCAGGCA 556
QY 122 yPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuProGl 142
Db 557 CGTCTCCGTTGGAGTACTGGACACGAGGAGGCGCTGGCTGGCAGAGGCGGCGGCGGCGGCGG 616
```



```
Db 562 C G T C T C C G T G G A G T A C T G G A C A C G G A G G C C T G G C C T G G C A G A G G C G C C G C C C A 621
Qy 142 u A s p P h e C y s P h e T y r G l n G l y S e r L e u A r g S e r H i s A r g A s n S e r S e r - - - V a l A l a l e 161
Db 622 C - - - - - T G C C T A G C G T G G T C A C C T G C A G G C C A G G C C A G C A G C T C C A T G T G G C C A T 675
Qy 161 u S e r T h r C y s G l n G l y L e u S e r G l y M e t I l e A r g T h r G l u G l u A l a A s p T y r P h e L e u A r 181
Db 676 C A G C A C C T G T G A G G C C T G C A G C C C T G A T C G T G C A G A C G A G A G A G T A C C T G A T T G A 735
Qy 181 g P r o L e u P r o S e r H i s L e u S e r T r p L y s L e u G l y A r g A l a A l a G l n G l y S e r S e r P r o S e 201
Db 736 G C C C C T G - - - - - C A G G T G G G C C C A A G G G T T C T G G A G C C C G A G A A A G T G G A C C A - - 787
Qy 201 r H i s V a l L e u T y r L y s A r g S e r T h r L y s L e u G l u P r o H i s A l a P r o G l y A l a S e r G l u V a l L e u V a 221
Db 788 - C A T G T G G T A C A G C G T T C C T C T - - - - - 811
Qy 221 l T h r S e r A r g T h r T r p G l u L e u A l a H i s G l n P r o L e u H i s S e r S e r A s p L e u A r g L e u G l 241
Db 812 - - J - - - - - C T G C T C A C C C C A C C T G - - - - - G A C A G C C T G T G G 843
Qy 241 Y l e u P r o G l n L y s G l n H i s P h e C y s G l y A r g - - - - - L y s L y s T y r M e t P r o G l n P r 259
Db 844 A G T G A G A G A T G A A A C C G T G A A A G G C G C C A T G T G C T G C G G A C C T T G A A G C C A C C 903
Qy 259 o P r o L y s G l u A s p L e u P h e I l e L e u P r o A s p G l u T y r L y s S e r C y s L e u A r g H i s L y s A r 279
Db 904 G C C T G C C A G A C C C T G G G A A T G A A C A G A G C G T G C C A G C G C T G - - - - - A A G C G 957
Qy 279 g S e r L e u L e u A r g S e r H i s A r g A s n G l u L e u A s n V a l G l u T h r L e u V a l V a l A s 299
Db 958 A T C G G T C A G C C G A - - - - - G A G C G T A C G T G G A G A C C C T G T G T G C T G A 1002
Qy 299 p l y s l y s M e t M e t G l n A s n H i s G l y H i s G l u A s n l e T h r T y r V a l L e u T h r I l e 319
Db 1003 C A A G A T G A T G T G G C C T A T C A C G G G C C G G G A T G T G G A G C A T A T G C T C C T G C C C A T C A T 1062
Qy 319 u A s n M e t V a l S e r A l a L e u P h e L y s A s p G l y T h r I l e G l y A s n I l e A s n I l e A l a I l 339
Db 1063 G A A C A T T G T G C C A A C T T T C C A G A C T C G A G T C T G G A A G C A C C G T T A A C A T C C T C G T 1122
Qy 339 e V a l G l y L e u I l e L e u L e u G l u A s p G l n P r o G l y L e u V a l l e S e r H i s H i s A l a A s 359
Db 1123 A A C T G C C C T A C T C T C A C G A G A C C A G C C C A C T C T G G A G A T C A C C C A C C A T G C G G 1182
Qy 359 p h i s T h r L e u S e r P h e C y s G l n T r p G l n S e r G l y L e u M e t G l y L y s A s p G l y - - - - - 377
Db 1183 G A A G T C C C T A G A C A G C T T C T G A A G T G G C A G A A T C C A T C G T G A C C A C A G C G C C A T G G 1242
Qy 378 - - - - - T h r A r g H i s A s p H i s A l a I l e L e u L e u T h r G l y L e 389
Db 1243 C A A T G C C A T C C A G A A C G T G T G C T A A C A T G A C A G C A G T C T C A T C A C A G C T A 1302
Qy 389 u A s p I l e C y s S e r T r p L y s A s n G l u P r o C y s A s p T h r L e u G l y P h e A l a P r o I l e S e r G l 409
Db 1303 T G A C A T C T C A T C T A C A A C A A C C C T G C G C A C A C T A G G C C T G G C C C C G G T G G C G G 1362
Qy 409 Y M e t C y s S e r L y s T y r A r g S e r C y s T h r I l e A s n G l u A s p T h r G l y L e u G l y L e u A l a P h 429
Db 1363 A A T G T G T A G C G C G A G A A G C T G C A G C G T C A A T G A G A C A T T G G C C T G C C A A A G C G T T 1422
Qy 429 e T h r I l e A l a H i s A s n P h e C l y H i s A s n P h e C l y M e t I l e H i s A s n G l y G l u G l y A s n M e 449
Db 1423 C A C C A T T G C C C A G A T C G G G C A C A C A T T C G G C A T G A A C C A T G A C G G C G T G G G A A A C A G 1482
Qy 449 t C y s - - - - - L y s L y s S e r G l u G l y A s n I l e M e t S e r P r o T h r L e u A l a G l y A r g A s 466
Db 1483 C T G T G G G C C C G T G T G C A G A C C C A G C C A A G C T A T G C T G C C C A C A T T A C C A T G A G A C 1542
Qy 466 n G l y V a l P h e S e r T r p S e r P r o C y s S e r A r g G l n T y r L e u H i s L y s P h e L e u A r g S e r 485
```

Db 1543 CAACCCATTGGTGGTGCATCTCTGCAACGGTGACTACATCACCAGCTTTCTAGACTCG 1600

Search completed: June 14, 2005, 16:00:23
Job time : 272 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2005, 11:25:01 ; Search time 714 Seconds
(without alignments)
4262.939 Million cell updates/sec

Title: US-10-804-457-2
Perfect score: 2628
Sequence: 1 MKPRARGWRGLAALMLLQA.....SPCSQYLHLFLRSVKMPAL 491

Scoring table:
BLASUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=slp
-O=/cgn2_1/USPTO spool_p/US10804457/runat_13062005_173511_3529/app query.fasta_1.647
-DB=Published Applications NA -QWMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blowsum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10804457 @CGN 1.1 723 @runat_13062005_173511_3529
-NCPU=6 -ICPU=3 -NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
5: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
6: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
8: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:
22: /cgn2_6/ptodata/2/pubpna/US10I_NEW_PUB.seq:
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2628	100.0	1476	13	US-10-217-774-1
2	2628	100.0	1476	21	US-10-804-457-1
3	2592	98.6	3675	13	US-10-217-774-3
4	2592	98.6	3675	15	US-10-296-616-1
5	2592	98.6	3675	21	US-10-804-457-3
6	2592	98.6	4042	13	US-10-217-774-5
7	2592	98.6	4042	21	US-10-804-457-5
8	2584	98.3	3675	17	US-10-330-176-1
9	2584	98.3	3675	18	US-10-275-107-16
10	1934	73.6	2997	10	US-09-981-151A-1
11	1929	73.4	2902	10	US-09-981-151A-1
12	1707	65.0	4888	17	US-10-399-645-19
13	1683	64.0	2895	10	US-09-981-151A-7
14	1298	49.4	3013	14	US-10-226-560-3
15	1298	49.4	3389	18	US-10-363-937-34
16	1298	49.4	3630	17	US-10-161-493-123
17	1297	49.4	2217	14	US-10-226-560-6
18	1295.5	49.3	1986	18	US-10-354-983-7
19	1295.5	49.3	3663	18	US-10-354-983-1
20	1293.5	49.2	2589	14	US-10-226-560-1
21	1293.5	49.2	3666	16	US-10-240-545A-1
22	1292.5	48.2	1530	14	US-10-226-560-4
23	1290	49.1	3555	17	US-10-188-186-129
24	1289.5	49.1	3219	18	US-10-354-983-3
25	1286	48.9	2940	17	US-10-188-186-131
26	1271.5	48.4	3445	19	US-10-391-364-86
27	1267	48.2	3246	19	US-10-391-364-88
28	1204	45.8	2433	10	US-09-981-151A-3
29	1156.5	45.0	3954	18	US-10-354-983-5
30	1039	39.5	2377	14	US-10-103-377C-1
31	1034.5	39.4	2052	14	US-10-103-377C-3
32	868	33.0	624	17	US-10-161-493-129
33	867	33.0	1059	17	US-10-161-493-129
34	864	32.9	1059	17	US-10-161-493-133
35	863	32.8	1059	17	US-10-161-493-125
36	863	32.8	1059	17	US-10-161-493-127
37	860	32.7	1059	17	US-10-161-493-131
38	825	31.4	2805	18	US-10-275-107-18
39	748	28.5	476	10	US-09-918-995-15335
40	721.5	27.5	5353	17	US-10-386-414-1
41	720.5	27.4	1407	9	US-09-963-791-5
42	720.5	27.4	1407	17	US-10-419-276-5
43	720	27.4	3179	17	US-10-172-118-1467
44	720	27.4	3179	18	US-10-342-887-1467
45	720	27.4	3179	21	US-10-956-157-1892

ALIGNMENTS

RESULT 1

US-10-217-774-1
; Sequence 1, Application US/10217774
; Publication No. US20020193583A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. US20020193583A1el Human Proteases and Polynucleotides Encoding
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/10/217,774
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US/09/930,872
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16


```
QY 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
Db 61 GTGGCCGAGCAGGACCTGGTGGCCATGGGACCCGCGAGCGCGCTGGAGCCCG 120
QY 41 SerValProArgProProProAlaGluArgProGlyTyrMetGluLysGlyGluTyr 60
Db 121 AGCGTCCCGCGTCTCTCCACCCGCGAGCGCGCGCTGGATGGAAAGGCGCAATAT 180
QY 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
Db 181 GACCTGGTCTCTGCGCTACGAGGTGGACCAAGGGCGGATTACGTGCCATGAATCATG 240
QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
Db 241 CACCATCAGCGCGGAGAGAGCAGTGGCGGTGTCGAGGTTGAGTCTCTTCACTTCGG 300
QY 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAla 120
Db 301 CTGAAAGGCTCCAGGCACGACTTCCACGCTGATCTGAGGACTTCCAGCAGCGCTAGTGGCT 360
QY 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 361 CTGGCTTTATTTGTCAGAGCTTGGGAAGACAGGCACTAAGTCTGTGAGACTTTACCG 420
QY 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db 421 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTCGATCACACAGAACTCTCTCAGTGGCC 480
QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu 180
Db 481 CTTTCACCTGCGCAGGCTTGTGAGCATGATACCAAGACAGAGCGAGATTACTTTCCTA 540
QY 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
Db 541 AGGCCACTTCTTACACCTCTCATGAACTCGCGAGCTGCCAGAGGAGCTCGCCA 600
QY 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
Db 601 TCCACGCTACTGTACAGAGATCCACAGAGCCCCATGCTCTGGGGCCAGTGAGGTCCTG 660
QY 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db 661 GTGACTCANAGACATGGAGCTGGACATCAACCCCTGCACAGAGGAGCTTCGCGCTG 720
QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
Db 721 GGACTGCCACAAAGCAGATTTCTGTGGAAGACGCAAGAAATACATGCCCGCCCTCCC 780
QY 261 LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db 781 AAGGAAGACCTCTTCATCTTCCAGATGAGTATAAGTCTTACGCGCATAGCGCTCT 840
QY 281 LeuLeuArgSerHisArgAsnGluLeuAsnValGluThrLeuValValValAspLys 300
Db 841 CTTCTGAGTCCCATAGAAATGAAAGTGAACGTGGAGACCTTGGTGGTGGTGCACAAA 900
QY 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrTyrValLeuThrIleLeuAsn 320
Db 901 AAGATGATGCAAAACCATGGCCATGAAATATCACCACTACGTGCTCAGCATACTCAAC 960
QY 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyAsnIleAsnIleAlaIleVal 340
Db 961 ATGGTATCTGCTTTATTCAAGATGGAAACATATGAGGAAACATCAATTTGCAATTGTA 1020
QY 341 GlyLeuIleLeuLeuGluAspGlnProGlyLeuValIleSerHisHisAlaAspHis 360
Db 1021 GGTCTGATTCTTCTAGAGATGAACAGCCAGGACTGTGTGATAGTCAACCCAGACACAC 1080
QY 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 1081 ACCTTAAGTAGTCTTCTGCGAGTGGAGTCTGGATTGAGGAAAGATGGAGCTCGTCAT 1140
QY 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
```

```
Db 1141 GACCAAGCCATCTTACTGACTGCTCGATATATGTTCTTGGAAAGATGAGCCCTGTGAC 1200
QY 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
Db 1201 ACTTTGGGATTGGACCCATTAAGTGAATGTGTAGTAAATATTCGAGCTGCACGATTAT 1260
QY 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 1261 GAAGATACAGGTCTTTGGGACTTCCCTTCCATTGCTCCATGAGTCTTGGACACAACTTTGGC 1320
QY 441 MetIleHisAspGlyGluGlyAsnMetCysValLysSerGluGlyAsnIleMetSerPro 460
Db 1321 ATGATTCTATGATGGGAAGGGAACATGTGTAAAGTCCGAGGCGCAATCATCATGTCCCT 1380
QY 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTrpLeuHis 480
Db 1381 ACATTGGCAGGACGCAATGGAGTCTTCTCTGGTCCACCTGCAGCGCCAGTATCTACAC 1440
QY 481 LysPheLeuArgSer-ValLysMetProAlaLeu 491
Db 1441 AAATTTCTAAGCAGCGCTCAAGCTATCTGCGCTTG 1474

RESULT 4
US-10-296-616-1
; Sequence 1, Application US/10296616
; Publication No. US20030129658A1
; GENERAL INFORMATION:
; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: No. US20030129658A1el protease
; FILE REFERENCE: Y0132PCT-664
; CURRENT APPLICATION NUMBER: US/10/296,616
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: JP 2000-393372
; PRIOR FILING DATE: 2000-12-25
; NUMBER OF SEQ ID NOS: 26
; SEQ ID NO 1
; LENGTH: 3675
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3675)
US-10-296-616-1

Alignment Scores:
Pred. No.: 3,17e-266 Length: 3675
Score: 2592.00 Matches: 486
Percent Similarity: 99.39% Conservative: 3
Best Local Similarity: 98.78% Mismatches: 2
Query Match: 98.63% Indels: 1
DB: 15 Gaps: 0

US-10-804-457-2 (1-491) x US-10-296-616-1 (1-3675)
QY 1 MetLysProArgAlaArgGlyTyrArgGlyLeuAlaAlaLeuTrpMetLeuLeuAlaGln 20
Db 1 ATGAGCCCGCGCGCGCGGATGGCGGGCTTGGCGGCGCTGTGGATGCTTTGGCGCGAG 60
QY 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
Db 61 GTGGCCGAGCAGGACCTGCTGCGCATGGGACCCGCGAGCGCGCTGGAGCCCG 120
QY 41 SerValProArgProProProAlaGluArgProGlyTyrMetGluLysGlyGluTyr 60
Db 121 AGCGTCCCGCGTCTCTCCACCCGCGAGCGCGCGCTGGATGGAAAGGCGCAATAT 180
QY 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
Db 181 GACCTGGTCTCTGCGCTACGAGGTGGACCAAGGGCGGATTACGTCTCCCATGAATCATG 240
QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
```

```
Db 241 CACCATCAGCGGCGGAGAGACAGTGGCGGTGTCGAGGTGCTCTTACACTTCGG 300
Qy LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerLeuValAla 120
Db 301 CTGAAGGGCTCCAGGCACGACTTCCACGTGGATCTCAGGACTTCCAGCAGCCTAGTGGCT 360
Qy ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 361 CCTGGCTTTATTGTGTGAGAGCTTTGGGAAAGACAGGCACTAAGTCTGTGAGACTTTACCG 420
Qy ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db 421 CCAGAGACTTCTGTTCTATCAGGCTCTTGGGATCAACAGAACTCTCTCAGTGGCC 480
Qy LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluAlaAspTyrPheLeu 180
Db 481 CTTTCAACCTGCCAAGGCTTGTCCAGCATGATACGAACAGAGAGCAGATTACTTCTTA 540
Qy ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaGlnGlySerSerPro 200
Db 541 AGGCCACTTCTTCCACACTCTCATGGAACTCGGCAGAGCTGCCCAAGGCAGCTGCCA 600
Qy SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
Db 601 TCCCAGTACTGTACAGAGATCCACAGAGCCCCATGCTCTGGGGCCAGTGAGGTCCTG 660
Qy ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db 661 GTGACCTCAAGGACATGGGAGCTGGGCACATCAACCTCTGCACAGCAGCAGCCTTCG 720
Qy GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
Db 721 GGACTGCCCAAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCCCAGCCTCC 780
Qy LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db 781 AAGGAAGACTCTTCACTTTCGAGATGAGTATAGTCTTGTCTTACGGCATAGCGCTCT 840
Qy LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValAspLys 300
Db 841 CTTCTGAGTCCCATAGAAATGAGAACTGAACGTGGAGACCTTGTGTGTGTCGACAA 900
Qy LysMetMetGlnAsnHisGlyHisGluAsnIleThrTyrValLeuThrIleLeuAsn 320
Db 901 AAGATGATGCMAAACCATGCCATGAATAATATCACCATCTAGTCTCAGCATCTCAAC 960
Qy MetValSerAlaLeuPheLysAspGlyThrIleGlyAsnIleAsnIleAlaIleVal 340
Db 961 ATGGTATCTGCTTTATTCAAGATGGAACAATAGGAGGAACATCAACATTGCAATTGA 1020
Qy GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis 360
Db 1021 GGTCTGATTTCTTAGAGATGAAACAGCCAGGACTGGTGATAGTCAACCCAGCAGACAC 1080
Qy ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 1081 ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCA 1140
Qy AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db 1141 GACCACGCCATCTTACTGACTGGTCTGGATATATGTTCTCGGAAGAAATGAGCCCTGTGAC 1200
Qy ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
Db 1201 ACTTTGGGATTTGCACCCATAGTGGAAATGTGTAGTAAATATCCAGCTGCAGATTAT 1260
Qy GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 1261 GAAGATACAGGTCTTGGACTGGCTTCCACATTGCCCATGAGTCTCGGACACAACTTTTGC 1320
Qy MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
Db 1321 ATGATTTCAATGAGAGAGGAAACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCT 1380
```

```
Qy 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
Db 1381 ACATTGGCAGGACGCAATGGAGTCTTCTCTGGTCACTCGCAGCGCGCATCTACAC 1440
Qy 481 LysPheLeuArgSerValLysMetProAlaLeu 491
Db 1441 AAATTTCTAAGCAGCGCTCAAGCTATCTGCCTTG 1474

RESULT 5
US-10-804-457-3
; Sequence 3, Application US/10804457
; Publication No. US20050065334A1
; GENERAL INFORMATION:
; APPLICANT: Hilbun, Carl Johan
; APPLICANT: Fziddle, Erin
; TITLE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/10/804,457
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: US/10/217,774
; PRIOR FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US/09/930,872
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3675
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-804-457-3
```

```
Alignment Scores:
Pred. No.: 3,178-266 Length: 3675
Score: 2592.00 Matches: 486
Percent Similarity: 99.33% Conservative: 3
Best Local Similarity: 98.78% Mismatches: 2
Query Match: 98.63% Indels: 1
DB: 21 Gaps: 0
```

US-10-804-457-2 (1-491) x US-10-804-457-3 (1-3675)

```
Qy 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaAlaLeuTrpMetLeuAlaGln 20
Db 1 ATGAAGCCCGCGCGCGGATGGCGGCTTGGCGGCGCTGTGGATGCTGTGGCGCAG 60
Qy 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
Db 61 GTGGCCGAGCAGCACCTGCGTGGCCATGGGACCCGCGCAGCGCGCTGGAGCCCCG 120
Qy 41 SerValProArgProProProProAlaGluArgProGlyTrpMetGluLysGlyGluTyr 60
Db 121 AGCGTCCCGCGTCTCTCTCCACCGCGGAGCGCGGCTGGATGGAAAGCGCGAATAT 180
Qy 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
Db 181 GACCTGGTCTCTGCGCTACGAGTTGACCAAGGGCGGATAGTGTCCCATGAAATCATG 240
Qy 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
Db 241 CACCATCAGCGCGGAGAGAGCAGTGGCGGTGTCGAGGTTGAGTCTCTTCCACTTCG 300
Qy 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerLeuValAla 120
Db 301 CTGAAGGGCTCCAGGCACGACTTCCACGTGGATCTCAGGACTTCCAGCAGCCTAGTGGCT 360
Qy 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 361 CCTGGCTTTATTGTGAGAGCTTTGGGAAAGACAGGCACTAAGTCTGTGAGACTTTACCG 420
```



```
QY 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
DB 699 TCCACGACTGTACAGAGATCCACAGAGCCCCATGCTCTCTGGGGCCAGTAGTGCTCTG 758
QY 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
DB 759 GTGACCTCAAGACATGGAGCTGGGCACATCAACCCCTGCACAGCAGCACCTTTGCCCTG 818
QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgLysLysTyrMetProGlnProPro 260
DB 819 CGACTGCCCAAAAGCAGCATTTCTGTGGAAGAGGCAAGAAATACATGCCCCAGCCTCCC 878
QY 261 LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
DB 879 AAGGAAGACCTCTTCATCTTGCAGATGAGTATAAGTCTTGCTTACGGCATTAAGCGCTCT 938
QY 281 LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValAspLys 300
DB 939 CTTCTGAGGTCCCATAGAAATGAGAACTGAACGTGGAGACCTTGGTGGTGGTGCANAA 998
QY 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrThrTyrValLeuThrIleLeuAsn 320
DB 999 AAGATGATGCAAAACCATGCCATGAAATATATCACCACCTACGTGCTCAGCATACTCAAC 1058
QY 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyAsnIleAsnIleAlaIleVal 340
DB 1059 ATGGTATCTGCTTTATTCAAGATGAAACAATAGGAGGAAACATCAACATTTGCAATTGTA 1118
QY 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis 360
DB 1119 GGTCTGATTCTTCTAGAAATGAACAGCCAGGACCTGGTGTAAAGTCAACCAAGACCAAC 1178
QY 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
DB 1179 ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGGAAGATGGGACTCGTCAAT 1238
QY 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
DB 1239 GACCACGCCATCTTACTGACTGCTCGGATATATGTTCTTGGAAAGATGAGCCCTGTGAC 1298
QY 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
DB 1299 ACTTTGGGATTTGCAACCAATAGTGAATGTGTAGTAAATATCGCAGCTGCACGATTAAT 1358
QY 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
DB 1359 GAAGATACAGGCTTGGACTGGCCCTCACCATTTGCCATGACTCTGGACACAACTTTGGC 1418
QY 441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
DB 1419 ATGATTATCATGGAGAGGAAACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCT 1478
QY 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrIleHis 480
DB 1479 ACATTGGCAGGACGCAATGAGTCTTCTCTGGTCACTCCCTGCAGCGCCAGTATCTACAC 1538
QY 481 LysPheLeuArgSerValLysMetProAlaLeu 491
DB 1539 AAATTTCTAAGCACCCGCTCAAGCTATCTGCCCTG 1572
```

RESULT 7

```
US-10-804-457-5
; Sequence 5, Application US/10804457
; Publication No. US2005006534A1
; GENERAL INFORMATION:
; APPLICANT: Fridde, Carl Johan
; TITLE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the
; TITLE OF INVENTION: Same
; FILE OF INVENTION: Same
; CURRENT APPLICATION NUMBER: US/10/804,457
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: US/10/217,774
```

```
; PRIOR FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US/09/930,872
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4042
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-804-457-5

Alignment Scores:
Pred. No.: 3,63e-266 Length: 4042
Score: 2592.00 Matches: 486
Percent Similarity: 99.3% Conservative: 3
Best Local Similarity: 98.7% Mismatches: 2
Query Match: 98.6% Indels: 1
DB: 21 Gaps: 0
```

US-10-804-457-2 (1-491) x US-10-804-457-5 (1-4042)

```
QY 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaAlaLeuTrpMetLeuLeuAlaGln 20
DB 99 ATGAAGCCCGCGCGCGCGATGGCGGGCTTTGGCGCGCTGTGGATGCTGTGGCGCAG 158
QY 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
DB 159 GTGGCGGAGCAGCAGCTGCGTGCCTATGGACCGCGCAGCGCGCTGGAGCCCG 218
QY 41 SerValProArgProProProAlaGluArgProGlyTrpMetGluLysGlyGluTyr 60
DB 219 AGCGTCCCGCGTCTCTCCACCCGCGAGCGCGCGCTGCGATGGAAAGGCGGAATAT 278
QY 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
DB 279 GACCTGGTCTCTGCGCTACGAGGTTGACCACAGGGCGGATACGTGTCCCATGAAATCATG 338
QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
DB 339 CACCATCAGCGCGGAGAGAGACAGTGGCGGTGTCGAGGTGAGTCTCTTACCTTCGG 398
QY 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAla 120
DB 399 CTGAAAGGCTCCAGCAGCAGCTTCCACGTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 458
QY 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
DB 459 CCTGGCTTTATTGTGCAGACGTTGGGAAAGACAGGCACCTAAGTCTGTGCAGACTTTACCG 518
QY 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
DB 519 CCAGAGGACTTCTGTGTTCTATCAGGCTCTTTGCCATCACACAGAAACTCTCTCAGTGGCC 578
QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu 180
DB 579 CTTTCAACTGCAAGGCTTGTACGAGCATGATACCAACAGAGAGGAGGAGGAGTACTTCTTA 638
QY 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
DB 639 AGGCCACTTCTCTCACACTCTCATGGAAGCTCGCGCAGAGCTGCCCAAGAGGAGCTGCCA 698
QY 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
DB 699 TCCACGACTACTGTACAAGAGATCCACAGAGCCCATGCTCTCTGGGGCCAGTGAGGCTCTG 758
QY 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
DB 759 GTGACCTCAGGACATGGAGCTGGCAGCATCAACCCCTGCACAGCAGCAGCCTTCGCTG 818
QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
```

```
Db      819  GGACTGCCACAAAGCAGCATTTCTGTGGAAAGCAGCAAGAAATACATGCCCCAGCCTCCC 878
QY      261  LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db      879  AAGGAAGACTCTTCACTCTTGCAGATAGTATAAGTCTTGCTTACGGCATAAAGCGCTCT 938
QY      281  LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValValAspLys 300
Db      939  CTTCTCAGGTCCCATAGAAATGAAGAACTGAAGCTGGAGACCTTGGTGGTGCACAAA 998
QY      301  LysMetMetGlnAsnHisGlyHisGluAsnIleThrThrTyrValLeuThrIleLeuAsn 320
Db      999  AAGATGATCAAAACCATGCCATGCAAAATATCACCACCTAGCTGCTCACGATACTCAAC 1058
QY      321  MetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAlaIleVal 340
Db      1059  ATGGTATCTGCTTTATTCANAGATGAACAATAGGAGGAAACATCAACATTGCAATTGTA 1118
QY      341  GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisHisAlaAspHis 360
Db      1119  GGTCTGATCTCTTAGAAGATGAACAGCCAGCAGCTGGTGAATAGTCACACGACCCAC 1178
QY      361  ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db      1179  ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCAT 1238
QY      381  AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db      1239  GACCAAGCCATCTTACTGACTGGTCTGGATATATGTTCTTGAAAGAAATGAGCCCTGTCAC 1298
QY      401  ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
Db      1299  ACTTTGGGATTTGCCACCATAGTGGAAATGTGTAGTAATATCGCAGCTGCACGATTAT 1358
QY      421  GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db      1359  GAAGATACAGGTCCTTGGACTGGCCCTTCACCATGTCCTGATGCTCGGACACAACTTTGGC 1418
QY      441  MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
Db      1419  ATGATTTCATGATGGAGAGGGNACATGTGTAAAGTTCGAGGGCAACATCATGTCCCT 1478
QY      461  ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
Db      1479  ACATTGGCAGGACGCAATGGAGTCTTCTCCTGGTCACCTGCAGCCGCGCAGTATCTACAC 1538
QY      481  LysPheLeuArgSer-ValLysMetProAlaLeu 491
Db      1539  AAATTTCTAAGCACCGCTCAAGCTATCTGCCTTG 1572
```

RESULT 8

```
US-10-330-176-1
; Sequence 1, Application US/10330176
; Publication No. US20030228676A1
; GENERAL INFORMATION:
; APPLICANT: Agostino, Michael
; APPLICANT: DiBlasio, Elizabeth
; TITLE OF INVENTION: AGGRECANASE MOLECULES
; FILE REFERENCE: AM100884
; CURRENT APPLICATION NUMBER: US/10/330,176
; CURRENT FILING DATE: 2002-12-30
; PRIOR APPLICATION NUMBER: 60/344,895
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 3675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-330-176-1
```

Alignment Scores:

Pred. No.: 2.28e-265 Length: 3675

```
Score: 2584.00 Matches: 484
Percent Similarity: 99.19% Conservative: 4
Best Local Similarity: 98.37% Mismatches: 3
Query Match: 98.33% Indels: 1
DB: 17 Gaps: 0

US-10-804-457-2 (1-491) x US-10-330-176-1 (1-3675)

QY      1  MetLysProArgAlaAraArgGlyTyrArgGlyLeuAlaLaLeuTrpMetLeuLeuAlaGln 20
Db      1  ATGAAGCCCGCGCGCGCGGATGGCGGGCTTGGCGGCGCTGTGGATGCTGCTGGCGCAG 60
QY      21  ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaAraProGlySerPro 40
Db      61  GTGGCCGAGCAGGACCTGCGTGGCCATGGACCCGCGGCGGCGGCTGGATGGAAAGGCGAATAT 120
QY      41  SerValProArgProProProAlaGluArgProGlyTyrMetGluLysGlyGluTyr 60
Db      121  AGCGTCCCGCGTCTCTCTCCACCCGCGGAGCGGCGGCTGGATGGAAAGGCGAATAT 180
QY      61  AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
Db      181  GACCTGGTCTCTGCGCTACGAGGTTGACCACAGGGCGGATTACGTGTCTCCCATGAATCATG 240
QY      81  HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
Db      241  CACCATCAGCGCGGAGAGAGCAGTGGCCGTGTCGAGGTTAGTCTCTTTCACCTTCGG 300
QY      101  LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAla 120
Db      301  CTGAAAGCCCCAGCAGCAGCTTCCACATGATCTGAGGACTTCCAGCAGCGCTAGTGCT 360
QY      121  ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db      361  CCTGCTTTATTTGTGCGACAGCTTGGGAAAGCAGGCACTAAGTCTGTGCGACACTTACCG 420
QY      141  ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db      421  CCAGAGACTTCTGTTTCTATCAAGGCTCTTTGGCATCACACAGAACTCTCTAGTGCC 480
QY      161  LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu 180
Db      481  CTTTCAACCTGCCAAGCTTGTCAAGGCATGTACAGAAAGAGGCGCAGATTAATCTTCTTA 540
QY      181  ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
Db      541  AGGCCACTTCTTCCACACCTCTCATGGAAACTCGGCGAGCTCGCCCAAGGCGAGCTCGCCA 600
QY      201  SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
Db      601  TCCCACTGACTGTACAGAGATCCACAGAGCCCCATGCTCTCTGGGCGCAGTGAGGCTCTG 660
QY      221  ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db      661  GTGACCTCAAGGACATGGGAGCTGGCACATCAACCCCTGCACAGCAGCAGCTTTCGCGCT 720
QY      241  GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
Db      721  GGACTGCCACAAAGAGCAGCATTTCTGTGGAAAGCGCAAGAAATACATGCCCGCCCTCCC 780
QY      261  LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db      781  AAGGAAGACTCTTTCATCTTGGCAGATGAGTATAAGTCTTGTCTTACGGCATTAAGCGTCT 840
QY      281  LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValValAspLys 300
Db      841  CTTCTGAGGTCCCATAGAAATGAAGAACTGAACGCTGGAGACCTTGTGTGTGGTGCACAAA 900
QY      301  LysMetMetGlnAsnHisGlyHisGluAsnIleThrThrTyrValLeuThrIleLeuAsn 320
Db      901  AAGATGATGCAAAACCATGGCCATGAAATATACCCACCTACGCTGCTCAGATACTCAAC 960
QY      321  MetValSerAlaLeuPheLysAspGlyThrIleGlyAsnIleAlaIleVal 340
```



```
Db 961 ATGGTATCTGCTTTATTTCAAGATGGAACAATAGGAGGAAACATCAACATTGCAATTGTA 1020
Qy 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis 360
Db 1021 GGTCTGATTCTTCTAGAAATGAACAGCCAGGACTGGTGTATAGTACACACGACAGACAC 1080
Qy 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyIleAspGlyThrArgHis 380
Db 1081 ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGGAAGATGGGACTCGTCAT 1140
Qy 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db 1141 GACCACGCCATCTTACTGACTGGTCTGGATATATGTTCTTGGAAAGATGAGCCCTGTGAC 1200
Qy 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTrpArgSerCysThrIleAsn 420
Db 1201 ACTTTGGGATTTCACCCATAGTGAATGTGTAGTAATATCGCAGCTGCAGATTAT 1260
Qy 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 1261 GAAGATACAGGCTCTGGACTGGCCCTTACCAATTGCCATGAGTCTGGACACAACTTTGGC 1320
Qy 441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
Db 1321 ATGATTTCATGATGGAGAGGGAACATGTGTAAAGTTCGAGGGCAACATCATGTGCCCT 1380
Qy 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTrpLeuHis 480
Db 1381 ACNTTGGCAGGAGCAGATGAGTCTCTCTGTCACCTTCGAGCGCCAGTATCTACAC 1440
Qy 481 LysPheLeuArgSerValLysMetProAlaLeu 491
Db 1441 AAATTTCTAAGCACCGCTCAAGCTATCTGCCTTG 1474
```

RESULT 9

```
US-10-275-107-16
; Sequence 16, Application US/10275107
; Publication No. US20040063107A1
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY D.
; APPLICANT: WHYTE, DAVID
; APPLICANT: SUDARSANAM, SUCHA
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEEL, SEAN R.
; APPLICANT: PAYNE, VILLA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1479
; CURRENT APPLICATION NUMBER: US/10/275,107
; CURRENT FILING DATE: 2003-11-03
; PRIOR APPLICATION NUMBER: PCT/US01/14431
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/201,879
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 3675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-275-107-16
```

```
Alignment Scores:
Pred. No.: 2,28e-265 Length: 3675
Score: 2584.00 Matches: 484
Percent Similarity: 99.19% Conservative: 4
Best Local Similarity: 98.37% Mismatches: 3
Query Match: 98.33% Indels: 1
DB: 18 Gaps: 0
```

US-10-804-457-2 (1-491) x US-10-275-107-16 (1-3675)

Qy 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaAlaLeuTrpMetLeuLeuAlaGln 20

```
Db 1 ATGAAGCCCGCGCGCGGATGGCGGGCTTTGGCGGCGCTGTGGATGCTGCTGGCGCAG 60
Qy 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaAlaProGlySerPro 40
Db 61 GTGGCCGAGCAGCAGCACCTGCGTGGCCATGGGACCGCAGCGCGCAGCGCTGGAGCCCG 120
Qy 41 SerValProArgProProProProAlaGluArgProGlyTrpMetGluLysGlyGluTrp 60
Db 121 AGCGTCCGCGCTCTCTCCACCCGCGAGCGCGCGGCTGGATGGAAAGCGCGCAATAT 180
Qy 61 AspLeuValSerAlaTrpGluValAspHisArgGlyAspTrpValSerHisGluIleMet 80
Db 181 GACCTGGTCTCTGCGCTACGAGGTTGACCACAGGGCGATTACGTGCTCCATGAATCATG 240
Qy 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
Db 241 CACCATCAGCGCGCGAGAGACAGTGGCGGTGTCGAGGTTGAGTCTCTTCCACCTTCGG 300
Qy 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAla 120
Db 301 CTGAAAGGCCCGCAGCAGCAGCTTCCACATGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360
Qy 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 361 CCTGCTTTTATTTGTGCAGACGTTGGGAAAGACAGGCACCTAAGTCTGTGCGACTTTTACC 420
Qy 141 ProGluAspPheCysPheTrpGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db 421 CCAGAGGACTTCTGTTTCTATCAAGCTCTTTTGGCATCACACAGAAACTCTCTCAGTGGCC 480
Qy 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTrpPheLeu 180
Db 481 CTTTCAACCTGCAAGGCTTGTTCAGGCATGTATACGAAACAGAGAGGAGCAGATTACTTCTTA 540
Qy 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
Db 541 AGGCCACTTCTCTTCCACCTCTCATGGAACTTCGGCAGAGCTGCCCAAGCAGCTCGCCA 600
Qy 201 SerHisValLeuTrpLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
Db 601 TCCACAGTACTGTACAGAGATCCACAGAGCCCCATGCTCTCTGGGGCCAGTGAGGTCTG 660
Qy 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db 661 GTGACCTCAAGGACATGGAGCTGCGACATCAACCCCTGCACAGCAGCAGCTTCGCTG 720
Qy 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTrpMetProGlnProPro 260
Db 721 GGACTGCCCAAAAGCAGCATTTCTGTGGAAGCAGCAAGAAATACATGCCCAGCCTCCC 780
Qy 261 LysGluAspLeuPheIleLeuProAspGluTrpLysSerCysLeuArgHisLysArgSer 280
Db 781 AAGGAGAGACTCTTTCATCTTGCAGATGAGTATAGTCTTGTCTACGGCATGAAGCGCTCT 840
Qy 281 LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValAspLys 300
Db 841 CTTCTGAGGTCCCATAGAAATGAAGAACTGAACGCTGGAGACCTTGGTGGTGGCGAATA 900
Qy 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrThrValLeuThrIleLeuAsn 320
Db 901 AAGATGATGCAAAACCATGGCCATGAAATATACCAACCTACGTGCTACGATACCTCAAC 960
Qy 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAlaIleVal 340
Db 961 ATGGTATCTGCTTTTATTCAGATGGACATAGGAGGAAACATCAACATTGCAATTGTA 1020
Qy 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis 360
Db 1021 GGTCTGATTCTTCTAGAGATGAACAGCCAGGACTGGTGTATAGTCAACCGCAGACAC 1080
Qy 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
```

```
Db 1081 ACCTAAGTAGCTTCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCAT 1140
Qy 381 AspHisAlaLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db 1141 GACACGCCATCTACTGACTGTGTGGATATATGTTCTCGGAAGATGAGCCCTGTGAC 1200
Qy 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTrpArgSerCysThrIleAsn 420
Db 1201 ACTTTGGATTGTCACCCATAAGTGGAAATGTGTAGTAATATGCCAGCTGCAGATTAT 1260
Qy 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 1261 GAAGATACAGGCTTGGACTGGCTTCCACCATGGCCATGAGTCTGGACACAACTTTTGC 1320
Qy 441 MetIleHisAspGlyGluGlyAsnMetCysLeuLysSerGluGlyAsnIleMetSerPro 460
Db 1321 ATGATTCATGATGAGAGAGGAACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCT 1380
Qy 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTrpLeuHis 480
Db 1381 ACATTGGCAGGACCAATGGAGTCTTCTCTGTGTCACCCCTGCAGCGCCAGTATCTACAC 1440
Qy 481 LysPheLeuArgSer-ValLysMetProAlaLeu 491
Db 1441 AATTTCTAAGCACCGCTCAAGCTATCTGCCTTG 1474

RESULT 10
US-09-981-151A-1
; Sequence 1. Application US/09981151A
; Publication No. US2003021256A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Gerlach, Valerie
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Muriel M
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: Feyman, John A
; APPLICANT: Stone, David J
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Shinkete, Richard A
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Gangolli, Esha R
; APPLICANT: Fernandes, Elma R
; APPLICANT: Gorman, Linda
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-168
; CURRENT APPLICATION NUMBER: US/09/981,151A
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/241,040
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,058
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,063
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,243
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/242,152
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/242,482
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,611
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,612
; PRIOR FILING DATE: 2000-10-23
```

```
; PRIOR APPLICATION NUMBER: 60/242,880
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,881
; PRIOR FILING DATE: 2000-10-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2997
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (857)..(858)
; OTHER INFORMATION: Wherein n is an a or t or c or g.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2383)
; OTHER INFORMATION: Wherein n is an a or t or c or g.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2983)
; OTHER INFORMATION: wherein n is an a or t or c or g.
US-09-981-151A-1

Alignment Scores:
Pred. No.: 5,99e-196 Length: 2997
Score: 1934.00 Matches: 402
Percent Similarity: 69.15% Conservative: 15
Best Local Similarity: 66.67% Mismatches: 43
Query Match: 73.59% Indels: 144
DB: 10 Gaps: 9

US-10-804-457-2 (1-491) x US-09-981-151A-1 (1-2997)
Qy 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaAlaLeuTrpMetLeuAlaGln 20
Db 10 ATGAAGCCCGCGCGCGCGGATGGCGGGGCTTGGCGGGCTGTGGATGCTGTGGCGGAG 69
Qy 21 ValAlaGlu----- 23
Db 70 GTGGCGGAGCAGGTGAGTCCCGGGCGTCCACACGCGCGGAAACCGCGGTCCGGACAG 129
Qy 23 ----- 23
Db 130 CTGGAGCGAGTCCCGCGCGGCTCTCTCCCGCGGACCCCGCGGTCTCACCGCGATGTCG 189
Qy 24 -----GlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySe 39
Db 190 CCGCTGTTTTCGCGAGGCACCTCGGTGGCGCATGGGACCCCGCAGCGCGCGCTGGGAG 249
Qy 39 rProSerValProArgProProProAlaGluArgProGlyTrpMetGlyLysGly 59
Db 250 CCCGAGCGTCCCGCGTCTCTCTCCACCGCGGAGCGCGCGGTGGATGAAAGGCGG 309
Qy 59 u----- 59
Db 310 -GACATGATGAAGCTGGAAACCATCGTCTCAGCAAACTAACACAGGAAACAAACCA 368
Qy 60 -----TyrAspLeuValSerAlaTyrGluValAspHisArgGly 72
Db 369 AACACTGCATGTTCTCACTCAATATGACCTGTCTCTGCTACGAGGTTGACCACAGGG 428
Qy 72 YAspTyrValSerHisGluIleMethHisGlnArgArgArgArgAlaValAlaValSe 92
Db 429 CGATTACGTGTCCCATGAATCATGACCATACGCGCGGAGAGAGAGTGTGCCCGTGC 488
Qy 92 rGluValGluSerLeuHisLeuArgLysGlySerArgHisAspPheHisValAspLe 112
Db 489 CGAGGTGTAGTCTCTTTCACCTTCGGCTGAAAGCCCCCGCAGCAGACTTCCACATGATCT 548
Qy 112 uArgThrSerSerSerLeuValAlaProGlyPheIleValGlnThrLeuGlyLysThrGly 132
```


Db 549 GAGGACTTCAGCAGCCTAGTGGCTCCTGGCTTTATTGTGCGACAGCTTGGGAAGCAGCAGG 608
Qy 132 YThrLysSerValGlnThrLeuProGluAspPheCysPheTyrGlnGlySerLeuAr 152
Db 609 CACTAAGTCTGTGCAGACTTACCGCCAGAGGACTTCTGTGTTCTATCAAGGCTCTTTTGG 668
Qy 152 gSerHisArgAsnSer---SerValAlaLeuSerThrCysGlnGlyLeuSerGlyMet11 171
Db 669 ATCACACAGAAATCGCCATCGCATGGAGGGAAGTTCTGTGAGGGCTCCACATCGCACTCT 728
Qy 171 eArg-----ThrGluGluAlaAspTyrPheLeuArgProLe 183
Db 729 GAAGCTCTGCAACAGTCAGAAATGTCCCGGGACAGTGTGACTTCGCTGCTGCTCAGTG 788
Qy 183 uProSerHisLeuSer-----TrpLys----- 190
Db 789 TGCCGAGCACACAGCAGCAGATTCAGAGGGCGGCACACTCAAGTGAAGCCTTACACTCA 848
Qy 190 ----- 190
Db 849 AGTAGAAGNNGACTTATGCNAACCTCTACTGTATCGCAGAAGGATTTGATTTCTTTTTC 908
Qy 191 -LeuGlyArgAlaAlaGlnGlySerSerProSerHisValLeuTyrLysArgSerThrGl 210
Db 909 TTGTCAAATAAAGTCAAAGATGGGACTCCATGCTCG-----GAGGATAGCCGTAA 959
Qy 210 uProHisAlaProGlyAlaSerGluValLeuValThrSerArgThrTrpGluLeuAlaHi 230
Db 960 TGTGTTGTATAGATGGATATGTGAGCTCAGTGTGGTGTCCACATCT-----CGCA 1010
Qy 230 sGlnProLeuHisSerAspLeuArgLeuGlyLeuProGlnLysGlnHisPheCysGl 250
Db 1011 C----- 1011
Qy 250 yArgArgLysLysTyrMetProGlnProGlnPheLeuAspLeuPheLeuLeuProAspGl 270
Db 1012 -----ATGCCCCAGCTCCCAAGGAAGACCTCTTCATCTTGCAGATGA 1055
Qy 270 uTyrLysSerCysLeuArgHisLysArgSerLeuLeuArgSerHisArgAsnGluGluLe 290
Db 1056 GTATAAGTCTTGCTTACGCATTAAGCGCTCTCTCTGAGTCCCATAGAAATGAAGACT 1115
Qy 290 uAenValGluThrLeuValValAspLysLysMetMetGlnAsnHisGlyHisGluAs 310
Db 1116 GAACTGGAGACCTTGGTGGTGTCACAAAAGATGATGCAAAACCATGGCCATGAAAA 1175
Qy 310 nileThrThrTyrValLeuThrIleLeuAsnMetValSerAlaLeuPheLysAspGlyTh 330
Db 1176 TATCACCTACCTAGTCTCAGCATCTCAGCATCTCAACATGGTATCTGCTTTATTCAAGATGGAAC 1235
Qy 330 rIleGlyGlyAsnIleAlaIleValGlyLeuIleLeuLeuGluAspGluGlnPr 350
Db 1236 ATAGAGGAGAAACATCAACATTGCAATTGAGTCTGATTCTCTTAGAAGATGAACAGCC 1295
Qy 350 oGlyLeuValIleSerHisHisAlaAspHisThrLeuSerSerPheCysGlnTrpGlnSe 370
Db 1296 AGSACTGGTGATAAGTCAACAGCAGCACACCTTAAGTAGCTTCTGCCAGTGGCAGTC 1355
Qy 370 rGlyLeuMetGlyLysAspGlyThrArgHisAspHisAlaIleLeuLeuThrGlyLeuAs 390
Db 1356 TGGATTGATGGGAAAGATGGGACTCGTCATGACACCGCATCTTACTGACTGGTCTGGA 1415
Qy 390 pIleCysSerTrpLysAsnGluProCysAspThrLeuGlyPheAlaProIleSerGlyWe 410
Db 1416 TATATGTTCTGGAGAGATGAGCCCTGTGACCTTTGGGATTTGGACCATTAAGTGAAT 1475
Qy 410 tCysSerLysTyrArgSerCysThrIleAenGluAspThrGlyLeuGlyLeuAlaPheTh 430
Db 1476 GTGTAGTAAATATCGCAGCTGCACGATTAATGAAGATACAGGTCTTGGACTGGCCCTTAC 1535
Qy 430 rIleAlaHisGluSerGlyHisAenPheGlyMetIleHisAspGlyGluGlyAsnMetCy 450
Db 1536 CATTTGCCCATGAGTCTGGACACAACTTTGGCATGATTCATGATGGAGAGGGAACATGTG 1595

Qy 450 sLysLysSerGluGlyAsnIleMetSerProThrLeuAlaGlyArgAsnGlyValPheSe 470
Db 1596 TAAAAGTCCGAGGGCAACATCATGTCCCTACATTGGCAGGACCAATGGAGTCTTCTC 1655
Qy 470 rTrpSerProCysSerArgGlnTyrLeuHisLysPheLeuArgSer-ValLysMetProA 490
Db 1656 CTGGTCACCTCGAGCGCCAGTATCTACACAAATTTTCTAAGCCCGCTCAAGCTATCTG 1715
Qy 490 laLeu 491
Db 1716 CCTTG 1720
RESULT 11
US-09-981-151A-5
; Sequence 5, Application US/09981151A
; Publication No. US20030212256A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Gerlach, Valerie
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Muriel M
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Feyman, John A
; APPLICANT: Stone, David J
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Shinkets, Richard A
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Gangolli, Bsha A
; APPLICANT: Fernandes, Elma R
; APPLICANT: Gorman, Linda
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-168
; CURRENT APPLICATION NUMBER: US/09/981,151A
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/241,040
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,058
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,063
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,243
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/242,152
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/242,482
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,611
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,612
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,880
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,881
; PRIOR FILING DATE: 2000-10-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5:
; LENGTH: 2902
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-981-151A-5

Alignment Scores:

Pred. No.: 1.96e-195 Length: 2902
 Score: 1929.00 Matches: 396
 Percent Similarity: 77.57% Conservative: 19
 Best Local Similarity: 74.02% Mismatches: 36
 Query Match: 73.40% Indels: 84
 DB: 10 Gaps: 8

US-10-804-457-2 (1-491) x US-09-981-151A-5 (1-2902)

QY 11 LeuAlaLeuAlaLeuTrpMetLeuLeuAlaGlnValAlaGlu----- 23
 Db 1 TTGGCGCGCTGTGGATGCTGCTGGCGAGTGGCGAGCGAGTGGCGCGCTCC 60
 QY 23 ----- 23
 Db 61 CACCAGCGGAAACCGCGGTCCGACAGCTGGAGCGAGTCCCGCGCTCTCTCC 120
 QY 24 -----GlnAlaProAlaCysAl 29
 Db 121 CGCGGACCCCGCGCTCTCACCGGATGTCGCGCTGTTTCCGCGACCTGGCGCG 180
 QY 29 aMetGlyProAlaAlaAlaProGlySerProSerValProArgProProAl 49
 Db 181 CATGGACCCGAGCGCGCGCTGGAGCGCGCGTCCGCGCTCTCTCTCCA----- 235
 QY 49 aGluArgProGlyTrpMetGluLysGlyGluTrpLeuValSerAlaTyGluValAs 69
 Db 236 -----GACCTGGTCTCTGCTACGAGGTGA 261
 QY 69 pHisArgGlyAspTyrValSerHisGluLeuMetHisGlnArgArgAlaVala 89
 Db 262 CACAGGGCGGATACGTCTCCCATGAATCATGACCATCAGCGCGGAGAGGAGT 321
 QY 89 lAlaValSerGluValGluSerHisGluArgLeuLysGlySerArgHisAspPhe 109
 Db 322 GCGCGTCCGAGTTGAGCCAGCTTCTCAGGTATGACGCGA-----GA 372
 QY 109 sValAspLeuArgThrSerSerSerLeuValAla-----ProGlyPheLeuValG 126
 Db 373 GCTCAGCTGTGTGGAGCGCTTCCCATGCTAATCTCAGCGCGGTTTGTGAACCT 432
 QY 126 nThrLeuGlyLysThr-----GlyThrLysSerValGlnThrLeuPr 140
 Db 433 TTCCAATGTCCTCTCACTGGAGGAAACAGCATGCTTCCAGAGAAATAAACA----- 487
 QY 140 oProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerValAl 160
 Db 488 -----AATGCAATGCTTGGAGATCGCCCTGCTTCAACCCAGAAAGTCTAATTGTGT 540
 QY 160 a-----LeuSerThrCysGlnGlyLeuSerGlyMetileArgThrGluGluAlaAspTy 178
 Db 541 TTTCTTTCTTTCTTTTATTTTTCAGTCAGGATGATACGAAACAGAGGAGGAGTGA 600
 QY 178 rPheLeuArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaGlnGlySe 198
 Db 601 CTCTTAAGGCGACTTCTTCTCACCTCTCATGAACTCGGAGAGTGGCGAGCGAG 660
 QY 198 rSerProSerHisValLeuTyLysArgSerThrGluProHisAlaProGlyAlaSerGl 218
 Db 661 CTGCGCATCCACGCTACTGTACAAGAGA-----GA 690
 QY 218 uValLeuValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLe 238
 Db 691 GGTCTGTGTGACCTCAAGGACATGGAGCTGGACATCAACCCCTGCACGACGCGACCT 750
 QY 238 uArgLeuGlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTrpMetProGl 258
 Db 751 TCGCGTGGAGTGCACAAAAGCAGCATTTCTGTGGAGAGCGCAGAAATACATGCCCCA 810
 QY 258 nProProLysGluAspLeuPheLeuLeuProAspGluTyLysSerCysLeuArgHisLy 278
 Db 811 GCCTCCCAAGGAGACCTCTTCTTCTTCCGAGATGATATAAGTCTTGTCTTACGCGATA 870

QY 278 sArgSerLeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValVa 298
 Db 871 GCGCTCTCTTCTGAGTCCCATAGAAATAGAAATGAAACGTGGAGACCTTGGTGGTGT 930
 QY 298 lAspLysLysMetMetGlnAsnHisGlyHisGluAsnIleThrThrValLeuThrIl 318
 Db 931 CGACAAAAGATGATGCAAAACCATATGATGCAAAATATCACCACCTAGTCTCAGCAT 990
 QY 318 eLeuAsnMetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAsnIleAl 338
 Db 991 ACTCAACATGGTATCTGCTTATTCAAAAGATGGAACAATAGAGGAAACATCAACATTC 1050
 QY 338 alLeValGlyLeuLeuLeuGluAspGluGlnProGlyLeuValIleSerHisHisAl 358
 Db 1051 AATTCTAGTCTGATTTCTTAGAAGATGAACAGCAGGACTGGTGATAGTCAACACGC 1110
 QY 358 aAspHisThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyTh 378
 Db 1111 AGACCACACCTTAAGTAGTTTCTGCCAGTGGCAGTCTGGATTGATGGGAAAGATGGAC 1170
 QY 378 rArgHisAspHisAlaLeuLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluPr 398
 Db 1171 TCGTCATGACCCAGCCCATCTTACTGACTGGTCTGGATATATGTTCTCTGGAAGATGAGCC 1230
 QY 398 oCysAspThrLeuGlyPheAlaProLysSerGlyMetCysSerLysTyArgSerCysTh 418
 Db 1231 CTGTGACACTTTGGGATTTTGACCCATAAGTGAATGTAGTAAATATATCGCAGCTGCAC 1290
 QY 418 rLeAsnGluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAs 438
 Db 1291 GATTAATGAAGATACAGTCTTGACCTGGCTTACCATTGCCCATGAGTCTGGACACAA 1350
 QY 438 nPheGlyMetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMe 458
 Db 1351 CTTTGGCATGATTCATGATGAGAGGAAACATGTGCAAAAGTCCCGAGGCAACATCAT 1410
 QY 458 tSerProThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTy 478
 Db 1411 GTCCCCCTACATTTGGCAGGAGCAATGAGTCTTCTCTGTGTCACCTTCGACCGCCAGTA 1470
 QY 478 rLeuHisPheLeuArgSer-VallysMetProAlaLeu 491
 Db 1471 TCTACACAAATTTCTAAGCACCCGCTCAAGCTATCTGCTTG 1511

RESULT 12

US-10-399-645-19
 ; Sequence 19, Application US/10399645
 ; Publication No. US20040029249A1
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE CORPORATION; LEE, Ernestine A.
 ; APPLICANT: HAFALIA, April J.A.; YUE, Henry
 ; APPLICANT: LAL, Preeti G.; YAO, Monique G.
 ; APPLICANT: LU, Yan; CHAWLA, Narinder K.
 ; APPLICANT: WARREN, Bridget A.; LU, Dying Aina M.
 ; APPLICANT: BAUGHN, Mariah R.; DELEGANE, Angelo M.
 ; APPLICANT: BURFORD, Neil; BOROWSKY, Mark L.
 ; APPLICANT: LEE, Sally; XU, Yuming
 ; APPLICANT: GRIFFIN, Jennifer A.; KALLICK, Deborah A.
 ; APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.
 ; APPLICANT: ISON, Craig H.; TANG, Y. Tom
 ; APPLICANT: AZIMZAI, Valda; ELLIOTT, Vicki S.
 ; APPLICANT: SWARNAKAR, Anita; RAMKUMAR, Jayalaxmi
 ; APPLICANT: NGUYEN, Dannie B.; TRIBOULEY, Catherine M.
 ; APPLICANT: LO, Terence P.; AU-YOUNG, Janice K.
 ; APPLICANT: THANGAVELU, Kavitha; KEARNEY, Liam
 ; TITLE OF INVENTION: PROTEASES
 ; FILE REFERENCE: PI-0263 USN
 ; CURRENT APPLICATION NUMBER: US/10/399,645
 ; CURRENT FILING DATE: 2003-04-16
 ; PRIOR APPLICATION NUMBER: PCT/US01/51034
 ; PRIOR FILING DATE: 2001-10-18
 ; PRIOR APPLICATION NUMBER: US 60/241,573

```
/ PRIOR FILING DATE: 2000-10-18
/ PRIOR APPLICATION NUMBER: US 60/243,643
/ PRIOR FILING DATE: 2000-10-25
/ PRIOR APPLICATION NUMBER: US 60/245,256
/ PRIOR FILING DATE: 2000-11-02
/ PRIOR APPLICATION NUMBER: US 60/248,395
/ PRIOR FILING DATE: 2000-11-13
/ PRIOR APPLICATION NUMBER: US 60/249,826
/ PRIOR FILING DATE: 2000-11-16
/ PRIOR APPLICATION NUMBER: US 60/252,303
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,981
/ PRIOR FILING DATE: 2000-12-01
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: PERL Program
/ SEQ ID NO 19
/ LENGTH: 4888
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. US20040029249A1 3533147CB1
US-10-399-645-19

Alignment Scores:
Pred. No.: 2,32e-171 Length: 4888
Score: 1707.00 Matches: 319
Percent Similarity: 99.08% Conservative: 3
Best Local Similarity: 98.15% Mismatches: 2
Query Match: 64.95% Indels: 1
DB: 17 Gaps: 0

US-10-804-457-2 (1-491) x US-10-399-645-19 (1-4888)

QY 168 SerGlyMetIleArgThrGluGluAlaAspTyrPheLeuArgProLeuSerHisLeu 187
DB 541 TCAGGCATATACGAACAGAGAGGAGATTAATCTCTAAGGCCATCTCTTCCACACCTC 600

QY 188 SerTrpLeuGlyArgAlaAlaGlnGlySerSerProSerHisValLeuTyrIleArg 207
DB 601 TCATGGAACTCGGCAGAGCTGCCAAGGAGCGCTGCCATCCACGCTACTGTACAAGAGA 660

QY 208 SerThrGluProHisAlaProGlyAlaSerGluValLeuValThrSerArgThrTrpGlu 227
DB 661 TCACAGAGCCCATGCTCTCGGGCCAGTGAGGTCTCTGACCTCAAGGACATGGGAG 720

QY 228 LeuAlaHisGlnProLeuHisSerSerAspLeuArgLeuGlyLeuProGlnHisGlnHis 247
DB 721 CTGGCACATCAACCCCTGCACAGCAGCGACCTTCGGCTGGGACTGCCACAAAGACGAT 780

QY 248 PheCysGlyArgGlyLysTyrMetProGlnProProLysGluAspLeuPheIleLeu 267
DB 781 TTCGTGGAGAGCGCAAGAAATACATGCCCGCCAGCTCCCAAGGAAGACCTTCTCATCTG 840

QY 268 ProAspGluTyrIleSerCysLeuArgHisLysArgSerLeuLeuArgSerHisArgAsn 287
DB 841 CCAGATGAGTATAGTCTTGTACGGCATAAGCGCTCTCTTCTGAGGTCCCATAGAAAT 900

QY 288 GluGluLeuAsnValGluThrLeuValValAspLysLysMetMetGlnAsnHisGly 307
DB 901 GAAGAACTGAACGTGGAGACCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 960

QY 308 HisGluAsnIleThrThrTyrValLeuThrIleLeuAsnMetValSerAlaLeuPheLys 327
DB 961 CATGAAATATATACCACTACGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTAC 1020

QY 328 AspGlyThrIleGlyAsnIleAsnIleAlaIleValGlyLeuLeuLeuGluAsp 347
DB 1021 GATGGAACAATAGGAGAAACATCAACATTCGAAATTTGAGTCTGATCTCTTAGAAGAT 1080

QY 348 GluGlnProGlyLeuValIleSerHisHisAlaAspHisThrLeuSerSerPheCysGln 367
DB 1081 GAACAGCCAGGACTGGTGTAGTAAAGTCAACCGCAGACACACCTTAAGTAGTCTTCTGCCAG 1140
```

```
QY 368 TrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHisAspHisAlaIleLeuLeuThr 387
DB 1141 TGGCAGTCTGGATTGATGGGGAAGATGGGACTCGTTCATGACCCACCCATCTTACTGACT 1200

QY 388 GlyLeuAspIleCysSerTrpLysAsnGluProCysAspThrLeuGlyPheAlaProIle 407
DB 1201 GGTCTGGATATATGTTCTCTGGAAGATGAGCCCTGTGACACTTTGGGANTTGCACCCATA 1260

QY 408 SerGlyMetCysSerLysTyrArgSerCysThrIleAsnGluAspThrGlyLeuGlyLeu 427
DB 1261 AGTGAATGTAGTAATATATGCGAGCTGCAAGATTAATGAAGATACAGTCTTGGACTG 1320

QY 428 AlaPheThrIleAlaHisGluSerGlyHisAsnPheGlyMetIleHisAspGlyGluGly 447
DB 1321 GCCTTCACCATTTGCCCATGAGTCTGGACACAACCTTTGGCATGATTCATGATGGAGAAGG 1380

QY 448 AsnMetCysLysLysSerGluGlyAsnIleMetSerProThrLeuAlaGlyArgAsnGly 467
DB 1381 AACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCTTACATTTGGCAGGACGCAATGGA 1440

QY 468 ValPheSerTrpSerProCysSerArgGlnTyrLeuHisLysPheLeuArgSer-Vally 487
DB 1441 GTCTTCTCTGTCACCTCGCCGCGCCAGTATCTACACAAATTTCTAAGCACCGCTCAA 1500

RESULT 13
US-09-981-151A-7
/ Sequence 7, Application US/09981151A
/ Publication No. US20030212256A1
/ GENERAL INFORMATION:
/ APPLICANT: Edinger, Shlomit R
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: MacDougall, John R
/ APPLICANT: Malyankar, Muriel M
/ APPLICANT: Smithson, Glennda
/ APPLICANT: Millet, Isabelle
/ APPLICANT: Peyman, John A
/ APPLICANT: Stone, David J
/ APPLICANT: Gunther, Erik
/ APPLICANT: Ellerman, Karen
/ APPLICANT: Shinkets, Richard A
/ APPLICANT: Padigaru, Muralidhara
/ APPLICANT: Guo, Xiaojia
/ APPLICANT: Patturajan, Meera
/ APPLICANT: Taupier Jr, Raymond J
/ APPLICANT: Burgess, Catherine E
/ APPLICANT: Zerhusen, Bryan D
/ APPLICANT: Kkuda, Ramesh
/ APPLICANT: Spytek, Kimberly A
/ APPLICANT: Gangolli, Esha A
/ APPLICANT: Fernandes, Elma R
/ APPLICANT: Gorman, Linda
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-168
/ CURRENT APPLICATION NUMBER: US/09/981,151A
/ CURRENT FILING DATE: 2001-10-16
/ PRIOR APPLICATION NUMBER: 60/241,040
/ PRIOR FILING DATE: 2000-10-17
/ PRIOR APPLICATION NUMBER: 60/241,058
/ PRIOR FILING DATE: 2000-10-17
/ PRIOR APPLICATION NUMBER: 60/241,063
/ PRIOR FILING DATE: 2000-10-17
/ PRIOR APPLICATION NUMBER: 60/241,243
/ PRIOR FILING DATE: 2000-10-17
/ PRIOR APPLICATION NUMBER: 60/242,152
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/242,482
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/242,611
```

; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/242,612
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/242,880
 ; PRIOR FILING DATE: 2000-10-24
 ; PRIOR APPLICATION NUMBER: 60/242,881
 ; PRIOR FILING DATE: 2000-10-24
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 160
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 7
 ; LENGTH: 2895
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-981-151A-7

Alignment Scores:
 Pred. No.: 4,08e-169 Length: 2895
 Score: 1683.00 Matches: 359
 Percent Similarity: 65.69% Conservative: 24
 Best Local Similarity: 61.58% Mismatches: 55
 Query Match: 64.04% Indels: 146
 DB: 10 Gaps: 11

US-10-804-457-2 (1-491) x US-09-981-151A-7 (1-2895)

QY 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaAlaLeuTrpMetLeuLeuAlaGln 20
 Db 10 ATGAAGCCCGCGCGCGGATGGCGGGCTTGGCGGCGCTGTGGATGCTGTGGCGGAG 69
 QY 21 ValAlaGlu----- 23
 Db 70 GTGGCCGAGCAGGTGAGTCCCGGCGCTCCACAGCGCGGAAACCGCGGGTCCGAGAC 129
 QY 23 ----- 23
 Db 130 CTGGAGCGAGTCCCGCGGCTCTCTCCCGGAGCCCGCGCTCACCGCGATGTCG 189
 QY 24 -----GlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySe 39
 Db 190 CGCGTGTGTTTCGCGCAGCACCTGCGTGGCCATGGGACCCGCGAGCGCGCTGGGAG 249
 QY 39 rProSerValProArgProProProAlaGluArgProGlyTrpMetGluLysGlyG 59
 Db 250 CCCGAGCGTCCCGCGCTCTCTCCACCGCGAGCGCGCGCTGGATGGAAAGGCGG 309
 QY 59 u----- 59
 Db 310 -GACATGGATGAAGCTGGAAACCATCGTTCTCAGCAAACTAACACAGGAACAGAAACCA 368
 QY 60 -----TyrAspLeuValSerAlaTyrGluValAspHisArgG 72
 Db 369 AACACTGCATGTTCTCACTCAATATGACCTGCTCTGCTACGAGGTTGACACAGGG 428
 QY 72 YAspTyrValSerHisGluLeuMetHisHisGluArgArgArgAlaValAlaValSe 92
 Db 429 CGATTACGTGTCCTCCATGAAATGTCACCATCAGCGCGGAGAGAGAGCGCGGTGTC 488
 QY 92 rGluValGluSerLeuHisLeuArgLeuLysGlySerArgHisAspPheHisValAspLe 112
 Db 489 CGAGGTTGAGTCTCTTCACTTCCGCTGAAAGGCGCCAGGCGACACTTCCACATGATCT 548
 QY 112 uArgThrSerSerLeuValAlaProGlyPheIleValGlnThrLeuGlyLysThrG 132
 Db 549 GAGACTTCCAGACCTTAGTGGCTCTCGGCTTTATTGTGCAGACGTTGGGAAGACAGG 608
 QY 132 YThrLysSerValGlnThrLeuProProGluAspPheCysPheTyrGlnGlySerLeuAr 152
 Db 609 CACTAAGTCTGTGCAGACTTTACCGCCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGCG 668
 QY 152 gSerHisArgAsnSer---SerValAlaLeuSerThrCysGlnGlyLeuSerGlyMetil 171
 Db 669 ATCACAGAAACTCGCCATCGCATGGAGGGAAGTTCTGTGAGGGCTCCCACTCGCACTCT 728

QY 171 eArg-----ThrGluGluAlaAspTyrPheLeuArgProle 183
 Db 729 GAAGCTCTGCAACAGCTCAGAAATGTCGCGGACAGTGTGACTTCCTCGTGCTGCTCAGTG 788
 QY 183 uProSerHisLeuSerTrpLysLeu---GlyArgAlaAlaGlnGlySerSerProSerHi 202
 Db 789 TCCCGAGCACAACAGCAGCAGATTCAGAGGCGG-----CA 824
 QY 202 sValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeuValTh 222
 Db 825 CTACAGTGGAGCCTTACACTCAAGTAGAAGCC---GACTTATGCAAACTCTACTGTAT 881
 QY 222 rSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeuGlyLe 242
 Db 882 CGCAGAAGGATTTGATTTCTTCTTTCT-----TTGTCAATAAAGTCAAGATGGGAC 935
 QY 242 uPro-----GlnLysGlnHisPheCys-----G 250
 Db 936 TCCATGCTCGGAGGATAGCCGTAATGTTGTATAGATGGGATATGTGAGCTCAGTGTGCT 995
 QY 250 YArgArgLysLysTyrMetProGlnProProLysGluAspLeuPheIleLeuProAspG 270
 Db 996 GTCCACATCTCGGCACATGCCCGAGCTCCACAGAGAGACCTCTCATCTTGCAGATGA 1055
 QY 270 uTyrLysSerCysLeuArgHisLysArgSerLeuLeuArgSerHisArgAsnGluGluLe 290
 Db 1056 GTATAAGTCTTGTCTACGGCATAAGCGCTCTCTCTGAGGTCCCATAGAAATGAAGAACT 1115
 QY 290 uAsnValGluThrLeuValValAspLysLysMetMetGlnAsnHisGlyHisGluAs 310
 Db 1116 GAACGTGGAGACCTTGTGTGGTGGCGACAAAAGATGATCAAAACCATGGCCCATGAAAA 1175
 QY 310 nIleThrThrTyrValLeuThrIleLeuAsnMetValSerAlaLeuPheLysAspGlyTh 330
 Db 1176 TATCACCACTTACGTGCTCAGGATACCTCAACATGTTATCTGCTTATTTCAAGAT----- 1230
 QY 330 rIleGlyGlyAsnIleAsnIleAlaIleValGlyLeuIleLeuLeuGluAspGluGlnPr 350
 Db 1230 ----- 1230
 QY 350 oGlyLeuValIleSerHisAlaAspHisThrLeuSerSerPheCysGlnTrpGlnSe 370
 Db 1230 ----- 1230
 QY 370 rGlyLeuMetGlyLysAspGlyThrArgHisAspHisAlaIleLeuLeuThrGlyLeuAs 390
 Db 1231 -GGAATTGATGGGAAAGATGGGACTCGTCTATGACACCGCATCTTACTGACTGCTGGA 1289
 QY 390 pIleCysSerTrpLysAsnGluProCysAspThrLeuGlyPheAlaProIleSerGlyMe 410
 Db 1290 TATATGTTCTGGAAGAAAGAGCCCTGTGACACTTTGGGATTTGCAACCCCAATAAGTGAAT 1349
 QY 410 tCysSerLysTyrArgSerCysThrIleAsnGluAspThrGlyLeuGlyLeuAlaPheTh 430
 Db 1350 GTGTAGTAAATATCGCAGCTGCACGATTAATGAAGATACAGGCTCTTGGACTGGCCTTAC 1409
 QY 430 rIleAlaHisGluSerGlyHisAsnPheGlyMetIleHisAspGlyGluGlyAsnMetCy 450
 Db 1410 CATTTGCCCATGAGTCTGGACACAACCTTTGGCATGATTATGATGGAGAGAGGGAACATGTG 1469
 QY 450 sLysLysSerGluGlyAsnIleMetSerProThrLeuAlaGlyArgAsnGlyValPheSe 470
 Db 1470 TAAAAAGTCCGAGGCAACATCATGTCTCCCTACATTTGGCAGGACCGCATTTGGAGTCTTCTC 1529
 QY 470 rTrpSerProCysSerArgGlnTyrIleuHisLysPheLeuArgSer-ValLysMetProA 490
 Db 1530 CTGGTCAACCTCGAGCGCCAGTATCTACACAAATTTCTTAAGCAGCGCTCAAGCTATCTG 1589
 QY 490 lAlaLeu 491
 Db 1590 CCTTG 1594

QY	395	sAenGluProCysAspThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyr	415
Db	1276	GAATGAACCATGTGACACTTAGGGTTTGCCCCCATCAGTGAATGTGCTTAAGTACCG	1335
QY	415	gSerCysThrIleAsnGluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSe	435
Db	1336	AAGTTGTACCATCAATGAGGACACAGGACTTGGCCTTGCCCTTACCATCGCTCATGAGTC	1395
QY	435	rGlyHisAsnPheGlyMetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGl	455
Db	1396	AGGGCACAACTTTGGTATGATTTCACGACGAGAGGGATCCCTGCAGAAAGGCTGAAGG	1455
QY	455	yAenIleMetSerProThrLeuAlaGlyArgAenGlyValPheSerTrpSerProCysSe	475
Db	1456	CAATATCATGTCTCCACACTGACCGGAAACAATGGAGTGTTCATGGTCTTCCTGCAG	1515
QY	475	rArgGlnTyrLeuHisLysPheLeu	483
Db	1516	CCGCCAGTATCTCAAGAAATTCCTC	1540

Search completed: June 14, 2005, 14:45:05
Job time : 754 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2005, 00:20:18 ; Search time 5228 Seconds
(without alignments)
4550.786 Million cell updates/sec

Title: US-10-804-457-2
Perfect score: 2628
Sequence: 1 MKPRARGWGLAALMLLAQ.....SPCSRQYLHKFLRSVKMPAL 491

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEVEL=xl
-Q=/cgn2_1/USPTO_spool_p/US10804457/runat_13062005_173509_3499/app_query.fasta_1.647
-DB=genEmbl -QFWT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=spect -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10804457@cgn 1.1 5600 @runat_13062005_173509_3499 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: gb.ba.*
- 2: gb.hcg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.ev.*
- 13: gb.un.*
- 14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2628	100.0	1476	6	AR229455	Sequence
2	2628	100.0	1476	6	AR534555	Sequence
3	2592	98.6	3675	6	BD170564	Novel
4	2592	98.6	3675	6	AR229456	Sequence

5	2592	98.6	3675	6	AR534556	Sequence
6	2592	98.6	4042	6	AR229457	Sequence
7	2592	98.6	4042	6	AR534557	Sequence
8	2587	98.4	3287	9	CQ850584	Sequence
9	2587	98.4	3287	9	AK127746	Homo sapi
10	2584	98.3	3227	9	HS315734	Homo sapi
11	2584	98.3	3675	6	AX139852	Sequence
12	2584	98.3	5610	6	BD170080	Novel
13	2584	98.3	8435	6	BD170083	Novel
14	2584	98.3	8505	6	BD170084	Novel
15	2324	88.4	1313	6	CQ738549	Sequence
16	2152	81.9	2225	9	AK122980	Homo sapi
17	2031.5	77.3	3402	10	BC034739	Mus muscu
18	2031.5	77.3	4904	10	AK173338	Mus muscu
19	1934	73.6	2997	6	AX574590	Sequence
20	1929	73.4	2902	6	AX574594	Sequence
21	1791.5	68.2	2166	6	AX763026	Sequence
22	1707	65.0	4888	6	AX786884	Sequence
23	1683	64.0	2895	6	AX574596	Sequence
24	1511	57.5	4234	6	BD170079	Novel
25	1511	57.5	4234	9	AB095949	Homo sapi
26	1300	49.5	5130	9	BC063283	Homo sapi
27	1298	49.4	3389	6	CQ819033	Sequence
28	1293.5	49.2	3666	6	BD170498	Novel
29	1287.5	49.0	4072	6	CQ845833	Sequence
30	1287.5	49.0	4072	9	AK131403	Homo sapi
31	1271.5	48.4	3445	9	HS311903	Sequence
32	1204	45.8	2433	6	AX574592	Sequence
33	1132	43.1	7668	6	BD170085	Sequence
34	1066	40.6	1668	6	AX481380	Sequence
35	1010	38.4	2172	6	AX799384	Sequence
36	963.5	36.7	2247	6	AX763025	Sequence
37	825	31.4	2805	6	AX319854	Sequence
38	721.5	27.5	5353	6	AX128433	Sequence
39	720.5	27.4	1407	6	AX430680	Sequence
40	720	27.4	3179	9	AX481383	Sequence
41	720	27.4	3179	9	AF140675	Homo sapi
42	718	27.3	5514	9	BC061631	Homo sapi
43	716.5	27.3	1524	6	AR430682	Sequence
44	716.5	27.3	1770	6	AR430683	Sequence
45	716.5	27.3	2727	6	AR430678	Sequence

ALIGNMENTS

RESULT 1
LOCUS AR229455 1476 bp DNA linear PAT 20-DSC-2002
DEFINITION Sequence 1 from patent US 6448388.
ACCESSION AR229455
VERSION AR229455.1 GI:27268949
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1476)
AUTHORS Fridde,C.J. and Hilbun,E.
TITLE Human proteases and polynucleotides encoding the same
JOURNAL Patent: US 6448388-A 1 10-SEP-2002;
FEATURES
source location/Qualifiers
1..1476
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:	1.14e-149	Length:	1476
Pred. No.:	2628.00	Matches:	491
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	100.00%	Gaps:	0
DB:	6		

US-10-804-457-2 (1-491) x AR229455 (1-1476)

QY 1 MetLysProArgAlaArgGlyTyrArgGlyLeuAlaLeuTrpMetLeuLeuAlaGln 20
Db 1 ATGAGCCCGCGCGCGGATGGCGGGCTTGGCGGCTGGATGCTGTTGGCGCAG 60
QY 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
Db 61 GTGGCCGAGCAGCAGCCTGCTGGCCATGGGACCGCGCAGCGGCTGGAGCCCG 120
QY 41 SerValProArgProProProAlaGluArgProGlyTyrMetGluLysGlyGluTyr 60
Db 121 AGCGTCCGCGTCTCTCTCCACCGCGGAGCGCGGCTGGATGAAAAGGCGGAATAT 180
QY 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
Db 181 GACCTGGTCTCTGCTTACGAGGTGGACCAAGGGCGGATAGTGTCCTCAAAATCATG 240
QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
Db 241 CACCATCAGCGCGGAGAGAGCAGTGGCGGTGCGAGGTTGAGTCTCTTCACCTTCG 300
QY 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAla 120
Db 301 CTGAAAGGCTCAGGACCGACCTTCCACGTGGATCTGAGGACTTCCAGCAGCCTAGTGCT 360
QY 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 361 CTGGGCTTTATTTGTGCAGAGTTGGGAAAGACAGGCATTAAGTCTGTGCAGACTTTACG 420
QY 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db 421 CCAGAGGACTTCTGTTTCTATCAGGCTCTTTGCGATCACAGAACTCTCTCAGTGCC 480
QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu 180
Db 481 CTTTCAACTGCGCAGGCTTGTGAGCATGATGAAACAGAGAGGCGAGATTACTTCTCA 540
QY 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
Db 541 AGGCCACTCTCTTACACTCTCATGGAACTCGGCAGAGTGGCCAGGAGCTCGCCA 600
QY 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
Db 601 TCCACAGTACTGTACAAGATCCACAGAGCCCATGCTCTCTGGGCGCAGTGAGGCTCTG 660
QY 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db 661 GTGACCTCAAGGACATGGAGCTGGACATCAACCCCTGCACAGCGGACCTTCGCGCTG 720
QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
Db 721 GGAGTGGCCACAAAAGCAGCATTTCTGTGGAGAGCGCAAGAAATACATGTCGCCAGCTCC 780
QY 261 LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db 781 AAGGAGACCTCTTCACTTCCAGATGAGTAAAGTCTTGTCTTACGGCATAAGCGCTCT 840
QY 281 LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValValAspLys 300
Db 841 CTTCTGAGGTCCCATAGAAATGAAGAACTGAACGTGGAGACCTTGGTGGTGGTCGACAA 900
QY 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrThrTyrValLeuThrIleLeuAsn 320
Db 901 AAGATGATGCAAAACCATGGCCATGAAAATATCACCACTACGTGCTTCAGTACTCAAC 960
QY 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyGluAsnIleAlaIleVal 340
Db 961 ATGGTATCTGCTTTATTCAAAGATGGAACTAATAGGAGAAACATCAACATTCGAATTGA 1020
QY 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisHisAlaAspHis 360
Db 1021 GGTCTGATCTTCTAGAAAGATGAACCCAGGACTGGTGATAGTCAACCCAGCAGACCAC 1080

QY 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 1081 ACCTTAAGTAGCTTCTGCCAGTGGCAGCTGGATTGATGGGGAAGATGGGACTGCTCAT 1140
QY 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db 1141 GACCACGCCCATCTTACTGACTGGTCTGGATATATGTTCTCTGGAGAATGAGCCCTGTGAC 1200
QY 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
Db 1201 ACTTTGGGATTGGACCCCATGAAGTGAATGTAGTAATAATATCGCAGCTGCACGATTAT 1260
QY 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 1261 GAAGATACAGGCTCTGGACTGGCTTACCATGTCCTCATGAGTCTGGACACACACTTTGGC 1320
QY 441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
Db 1321 ATGATTCTATGATGGAGAGGGAACATGTGTAAAAGTCCGAGGCAACATCATGTGCCCT 1380
QY 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
Db 1381 ACATTGGCAGGACCAATGGAGTCTTCTCTGGTCACTTCACCTGCAGCCGCCAGTATCTAC 1440
QY 481 LysPheLeuArgSerValLysMetProAlaLeu 491
Db 1441 AAATTTCTAAGATCAGTGAAATGCCAGCTCTC 1473
RESULT 2
AR534555
LOCUS AR534555 1476 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6734007.
ACCESSION AR534555
VERSION AR534555.1 GI:53924880
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1476)
AUTHORS Fridde,C.J. and Hilbun,E.
TITLE Human proteases and polynucleotides encoding the same
JOURNAL Patent: US 6734007-A 11-MAY-2004;
FEATURES
source
1. .1476
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1.14e-149 Length: 1476
Score: 2628.00 Matches: 491
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-804-457-2 (1-491) x AR534555 (1-1476)
QY 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaAlaLeuTrpMetLeuLeuAlaGln 20
Db 1 ATGAGCCCGCGCGCGGATGGCGGGCTTGGCGGCTGGATGCTGTTGGCGCAG 60
QY 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
Db 61 GTGGCCGAGCAGCAGCCTGCTGGCCATGGGACCGCGCAGCGGCTGGAGCCCG 120
QY 41 SerValProArgProProProAlaGluArgProGlyTyrMetGluLysGlyGluTyr 60
Db 121 AGCGTCCGCGTCTCTCTCCACCGCGGAGCGCGGCTGGATGAAAAGGCGGAATAT 180
QY 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
Db 181 GACCTGGTCTCTGCTTACGAGGTGGACCAAGGGCGGATAGTGTCCTCAAAATCATG 240

```
Db 181 GACCTGGTCTCTGCTACGAGTTGACCAACAGGGGGCGATTACGTGTCCCATGAATCATG 240
Qy 81 HisHisGlnArgArgArgAlaValaValSerGluValGluSerLeuHisLeuArg 100
Db 241 CACCATCAGCGCGGAGAGAGCAGTGGCCGTGTCGAGGTTGAGTCTCTTCACCTTCGG 300
Qy 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerLeuValAla 120
Db 301 CTGAAGAGCTCCAGGCGACGACTTCACGTGGATCTGAGGACTTCAGCAGCGCTAGTGCT 360
Qy 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 361 CCTGGCTTATGTGCAGACGTTGGGAAGACAGCCACTAAGTCTGTGCAGACTTTACCG 420
Qy 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db 421 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGGATCACACAGAAACTCTCAGTGCC 480
Qy 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu 180
Db 481 CTTTCAACCTGCCAAGGCTTCTCAGGCATGTACGAACAGAGAGGCGCAGATTACTTCTTA 540
Qy 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
Db 541 AGGCCACTTCTTCACACTCTCATGGAACTCCGCGAGCTGCCAAGCGAGCTCGCCA 600
Qy 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
Db 601 TCCCACTGACTGTACAGAGATCCACAGAGCCCCATGCTCTGGGGCCAGTGAGTCTCG 660
Qy 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db 661 GTGACCTCAGAGACATGGAGCTGGCACATCAACCCCTGTCACAGCAGCGACTTCGCGTG 720
Qy 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
Db 721 GGACTGCCACAAAGAGCAGACTTCTGTGGAAGACGCAAGAAATACATGCCCGCCCTCCC 780
Qy 261 LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db 781 AAGGAGACCTCTTCATCTTCCAGATGAGTATAGTCTTGCTTACGGCATAGCGCTCT 840
Qy 281 LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValAspLys 300
Db 841 CTTCTGAGGTCCCATAGAAATGAAGAACTGAACGTGGAGACTTGTGGTGGTGCACAAA 900
Qy 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrThrTyrValLeuThrIleLeuAsn 320
Db 901 AAGATGATGCAAAACCCATGGCCATGAATAATACCACTACGTGCTCACGATACTCAAC 960
Qy 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAsnIleAlaIleVal 340
Db 961 ATGTATCTGTTTATTCNAAGATGGAAACATAGAGGAAACATCAACATTCGATTTGTA 1020
Qy 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisHisAlaAspHis 360
Db 1021 GGTCTGATTCTCTAGAAGATGAACAGCAGCACTGGTGATAAGTCAACACGACGAGCCAC 1080
Qy 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 1081 ACCTTAAGTAGTCTTCTGCGAGTGGCAGTCTGGATTTGATGGGGAAGATGGGACTCGTCAT 1140
Qy 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db 1141 GACCACGCCATCTTACTGACTGGTCTGGATATATGTTCTCTGGAAAGATGAGCCCTGTGAC 1200
Qy 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
Db 1201 ACTTTGGGATTTGCACCCATAAGTGAATGTGTAGTAATAATATCGCAGCTGCACGATTAAT 1260
Qy 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 1261 GAAGATACAGGTCTTGGACTGGCCTTCACCATGGCCATGAGTCTGGGACACAACTTTGGC 1320
```

```
Qy 441 MetIleHisAspGlyGlyGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
Db 1321 ATGATTTCATGTAGAGAGGGAACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCT 1380
Qy 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
Db 1381 ACATGGCAGACCAATGGAGTCTTCTCTGGTGCACCTGCAGCCGCAGTATCTACAC 1440
Qy 481 LysPheLeuArgSerValLysMetProAlaLeu 491
Db 1441 AAATTTCTAAGATCAGTGAAATGCCAGTCTC 1473

RESULT 3
BD170564
LOCUS Novel protease. 3675 bp DNA linear PAT 17-JAN-2003
ACCESSION BD170564
VERSION BD170564.1 GI:27876376
KEYWORDS WO 02051998-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3675)
AUTHORS Yamaji,N., Nishimura,K., Abe,K. and Ogino,M.
TITLE Novel protease
JOURNAL Patent: WO 02051998-A 1 04-JUL-2002;
YAMANOUCHI PHARMACEUTICAL CO LTD,NOBORU YAMAJI, KOICHI NISHIMURA,
KUNITAKE ABE, MAKOTO OGINO
COMMENT OS Homo sapiens (human)
PN WO 02051998-A/1
PD 04-JUL-2002
PF 21-DEC-2001 WO 2001JP011251
PR 25-DEC-2000 JP 00P 393372
PI NOBORU YAMAJI, KOICHI NISHIMURA, KUNITAKE ABE, MAKOTO OGINO PC
C12N15/09, C07K16/40, C12N9/10, C12N33/15, G01N33/50, PC
G01N33/573//

FEATURES
source
1..3675
Location/Qualifiers
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 5,09e-147 Length: 3675
Score: 2592.00 Matches: 486
Percent Similarity: 99.39% Conservative: 3
Best Local Similarity: 98.78% Mismatches: 2
Query Match: 98.63% Indels: 1
DB: 6 Gaps: 0

US-10-804-457-2 (1-491) x BD170564 (1-3675)
Qy 1 MetLysProArgAlaArgGlyTyrArgGlyLeuAlaLeuTrpMetLeuLeuAlaGln 20
Db 1 ATGAAGCCCGCGCGCGGATGGCGGCTGTGGCGGCTGTGGATGCTGTGGCGCAG 60
Qy 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
Db 61 GTGGCCGAGCAGGACCTCGCGCATGGGACCCGAGCGGCGGCTGTGGAGCCCG 120
Qy 41 SerValProArgProProProAlaGluArgProGlyTrpMetGluLysGlyGluTyr 60
Db 121 AGCGTCCGGGCTCTCTCCACCCCGCGAGCGGCGCGGCTGTGGATGGAAAGGCGAATAT 180
Qy 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
```



```
Db 421 CAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGCATCACAGAAATCTCTCATGTGGCC 480
Qy 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluAlaAspTyrPheLeu 180
Db 481 CTTTCAACTGCAAGGCTTGTGAGCATGATACGAAGAAGAGCGAGATTACTTCTCTA 540
Qy 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
Db 541 AGGCCACTTCTTTCACACCTCTCATGGAATCTCGGACAGCTGCCCAAGCGAGCTCGCCA 600
Qy 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyValAsnGluValLeu 220
Db 601 TCCACAGTACTGTACAAGATCCACAGACCCCATGCTCTCTGGGGCCAGTGAGTCTCTG 660
Qy 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db 661 GTGACCTCAAGGACATGGAGCTGGACATCAACCCCTGCACAGCAGCGACCTTCGCTG 720
Qy 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
Db 721 GGACTGCCACAAAGCAGCATTTCTGTGGAGACGCAAGAAATACATGCCCGAGCTGCC 780
Qy 261 LysGluAspLeuPheIleuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db 781 AAGGAAGACCTCTTCATCTTGCCAGATGAGTATAAGTCTTGTCTACGGCATAGCGCTCT 840
Qy 281 LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValAspLys 300
Db 841 CTTCTGAGGTCCATAGAAATGAAGACTGAACGTGGAGACCTTGGTGGTGCACAAA 900
Qy 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrTyrValLeuThrIleLeuAsn 320
Db 901 AAGATGATGCAAAACCATGGCCATGAAATATACCACTAGCTCTCACCATATCTCAAC 960
Qy 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyLysAsnIleAsnIleAlaVal 340
Db 961 ATGGTATCTGCTTTTATCAAGATGAAACATAGAGGAAACATCAACATTGCAATTGTA 1020
Qy 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis 360
Db 1021 GGTCTGATTCTTCTAGAGATGAACAGCAGACCTGGTGATAGTACCACGACGACAC 1080
Qy 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 1081 ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCA 1140
Qy 381 AspHisAlaIleLeuLeuThrClyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db 1141 GACCAAGCCATCTTACTGACTGGTCTGGATATATGTTCTTGGAAAGATGAGCCCTGTGAC 1200
Qy 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
Db 1201 ACTTTGGGATTGTGACCCCATAGTGGATGTGTAGTAAATATCGCAGCTGCACGATTAT 1260
Qy 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 1261 GAAATACAGGCTTGTGACATGGCCCTTCAACATGGCCATGAGTCTGGACACAACTTTGGC 1320
Qy 441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
Db 1321 ATGATTTCATGATGGAGAAGGGAACATGTGTAAAAAGTCCGAGGGCAACATCATGTGCC 1380
Qy 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
Db 1381 ACATTGGCAGGACGCAATGAGTCTTCTCTGGTCACTCCCTGACCGCGCCAGTATCTACAC 1440
Qy 481 LysPheLeuArgSer-ValLysMetProAlaLeu 491
Db 1441 AAATTCTTAAGCACCGCTCAAGCTATCTGCCCTTG 1474
```

RESULT 5
AR534556
LOCUS

DEFINITION Sequence 3 from patent US 6734007.
ACCESSION AR534556
VERSION AR534556.1 GI:53924881
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3675)
AUTHORS Fridde,C.J. and Hilbun,E.
TITLE Human proteases and polynucleotides encoding the same
JOURNAL Patent: US 6734007-A 3 11-MAY-2004;
FEATURES Location/Qualifiers
source 1..3675
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 5,09e-147 Length: 3675
Score: 2592.00 Matches: 486
Percent Similarity: 99.39% Conservative: 3
Best Local Similarity: 98.78% Mismatches: 2
Query Match: 98.63% Indels: 1
DB: 6 Gaps: 0

US-10-804-457-2 (1-491) x AR534556 (1-3675)

```
Qy 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaAlaLeuTrpMetLeuLeuAlaGln 20
Db 1 ATGAAGCCCGCGCGCGGATGGCGGGCTTGGCGGGCTGTGGATGTGTGTGGCGAG 60
Qy 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaAlaProGlySerPro 40
Db 61 GTGGCCGAGCAGGACCTGCGTGCCTATGGACCCGAGCGGAGCGCTGGAGCGCG 120
Qy 41 SerValProArgProProProProAlaGluArgProGlyTrpMetGluLysGlyLeuTyr 60
Db 121 AGCGTCCCGCTCTCTCCACCCCGGAGCGCGGCTGGATGGTGAAGGCGGAATAT 180
Qy 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluMet 80
Db 181 GACCTGTCTCTGCTCAGAGTTGACCAAGGCGGATAGCTGTCCATGAAATCATG 240
Qy 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
Db 241 CACCATCAGCGCGGAGAGAGAGAGTGGCGCTGTCCGAGGTTGAGTCTCTTCACTTCGG 300
Qy 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAla 120
Db 301 CTGAAAGGCTTCAGGACGACCTTCCACGTGGATCTTGAGACTTCCAGCAGCCTAGTG 360
Qy 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 361 CTGTGCTTTATTGTGCAGACGTTGGGAAAGACAGACACTAAGTCTGTGCAGACTTTAC 420
Qy 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db 421 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGCATCACACAGAAACTCTCTCAGTGG 480
Qy 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu 180
Db 481 CTTTCAACTGCAAGGCTTGTGAGCATGATACGAAGAAGAGCGAGATTACTTCTCTA 540
Qy 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
Db 541 AGGCCACTTCTTTCACACCTCTCATGGAATCTCGGACAGCTGCCCAAGCGAGCTCGCCA 600
Qy 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyValAsnGluValLeu 220
Db 601 TCCACAGTACTGTACAAGATCCACAGACCCCATGCTCTCTGGGGCCAGTGAGTCTCTG 660
Qy 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
```

```
Db 661 GTGACCTCAAGGACATGGGAGCTGGCACAATCAACCCCTGCACAGCGACCTTGGCTG 720
QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
Db 721 GGATTCGCCAAGACAGCATTTCTGTGGAGAGCGAAGAAATACATGCCCGACCTCC 780
QY 261 LysGluAspLeuPheIleLeuProAspGlnTyrLysSerCysLeuArgHisLysArgSer 280
Db 781 AAGGAAGACCTCTTCACTTCCAGAGATGATTAAGTCTTTCAGGCATAGAGCGCTCT 840
QY 281 LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValAspLys 300
Db 841 CTTCTGAGGTCCTCATGAAATGAGAACTGAAGCTGGAGACCTTGGTGGTGGCGA 900
QY 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrTyrValLeuThrIleLeuAsn 320
Db 901 AAGATGATGCMAAACCATGSCCATGAAATATCACCACCTACGCTCAGTACTCAAC 960
QY 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyAsnIleAlaIleVal 340
Db 961 ATGGTATCTGCTTTATTCAAGATGAAACAATAGGAGGAAACATCAACATTGCATTTGA 1020
QY 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis 360
Db 1021 GGTCTGATCTTCTAGAGATGAACAGCCAGGACTGGTGAATAGTCAAGTCAACAGCACC 1080
QY 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 1081 ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCA 1140
QY 381 AspHisAlaIleLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db 1141 GACCAAGCCATCTTACTGACTGGTCTGGATATATGTTCTGGAGAAATGAGCCCTGTGAC 1200
QY 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
Db 1201 ACTTTGGGATTTGCCCATAGTGGATGTGTAGTAATATCGCAGCTGCACGATTAAT 1260
QY 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 1261 GAAGATACAGGTCCTTGGACTGGCTTCCATCCATGCTGCCATGAGTCTGGACACAACTTTGGC 1320
QY 441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
Db 1321 ATGATTTCATGATGGAGAGGAGACATGTGTAAAGTCCGAGAGGCAACATCATGTCCCT 1380
QY 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
Db 1381 ACATTGGCAGGACGCAATGGAGTCTTCTCTGCTCACCTGCACCCCGCCAGTATCTACAC 1440
QY 481 LysPheLeuArgSerValLysMetProAlaLeu 491
Db 1441 AAATTTCTAAGCAGCGCTCAAGCTATCTGGCTTG 1474

RESULT 6
LOCUS AR229457
DEFINITION Sequence 5 from patent US 6448398.
ACCESSION AR229457
VERSION AR229457.1 GI:27268951
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE
1 (bases 1 to 4042)
AUTHORS Friddle,C.J. and Hilbun,E.
TITLE Human proteases and polynucleotides encoding the same
JOURNAL Patent: US 6448398-A 5 10-SEP-2002;
FEATURES
Location/Qualifiers
source 1..4042
/organism="unknown"
/mol_type="genomic DNA"
```

Alignment Scores:

Pred. No.: 5,719-147 Length: 4042
Score: 2592.00 Matches: 486
Percent Similarity: 99.39% Conservative: 3
Best Local Similarity: 98.78% Mismatches: 2
Query Match: 98.63% Indels: 1
DB: 6 Gaps: 0

US-10-804-457-2 (1-491) x AR229457 (1-4042)

```
QY 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaLeuTrpMetLeuLeuAlaGln 20
Db 99 ATGAAGCCCGCGCGCGGATGCGGGGCTTGGGGCGCTGTGGATGCTGCTGGCGGAG 158
QY 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
Db 159 GTGCCGAGCAGGACCTTGGCGTGGCCATGGGACCCGACGCGGCGCGCTGGGAGCCG 218
QY 41 SerValProArgProProProAlaGluArgProGlyTrpMetGluLysGlyGluTyr 60
Db 219 AGCGTCCCGGCTCTCTCTCCACCGCGGAGCGCGGCTGGATGGAAAGGGCGAATAT 278
QY 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
Db 279 GACCTGTCTCTGCTTACGAGGTTGACCAAGGGGCGATTAACGTGCCATGAATCATG 338
QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
Db 339 CACCATCAGCGCGGAGAGAGCAGTGGCCGTGTCCAGGTTGAGTCTCTTCACCTTCGG 398
QY 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAla 120
Db 399 CTGAAGAGCTTCCAGGACGACTTCCACGTCGATCTCAGGACTTCCAGCAGCTAGTGGCT 458
QY 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 459 CCTGGCTTATTGTGCAGAGCTTGGGAAAGACAGGCACCTAAGTCTGTGCAGACTTTACCG 518
QY 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db 519 CCAGAGGACTTCTGTCTTCTATCAAGGCTCTTTGGCATCACACAGAAACTCTCAGTGGCC 578
QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu 180
Db 579 CTTTCAACTGCAAGGCTTGTCCAGCATGATACGAACAGAGAGCGCAGATTACTCTTA 638
QY 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerPro 200
Db 639 AGGCCACTTCTTTCACCTCTCATGAAACTCGCGCAGAGCTGCCAAGGAGCTCGCCA 698
QY 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyValAspGluValLeu 220
Db 699 TCCACGCTACTGTACAGAGATCCACAGAGCCCATGCTCTCTGGGGCCAGTGAGTCTCTG 758
QY 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db 759 GTGACCTCAAGGACATGGAGCTGGCACAATCAACCCCTGCACAGAGCGACTTTCGCTG 818
QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
Db 819 GGACTGCCCAAAAGCAGCATTTCTGTGGAGAGCGAAGAAATACATGCCCGCCAGCTCCC 878
QY 261 LysGluAspLeuPheIleLeuProAspGlnTyrLysSerCysLeuArgHisLysArgSer 280
Db 879 AAGGAAGACCTCTTCACTTCTTGGCAGATGATTAAGTCTTTCGCTTACGGCATTAAGCGCTCT 938
QY 281 LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValAspLys 300
Db 939 CTTCTGAGGTCCTCATAGAAATGAAGAACTGAACGTGGAGACCTTGGTGGTGGTCGACAA 998
QY 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrThrTyrValLeuThrIleLeuAsn 320
```



```
Db 999 AAGATGATGCAAAACCATGGCCATGAATAATACACCACCTAGCTGCTACGATCTCAAC 1058
Qy 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAlaIleVal 340
Db 1059 ATGGTATCTGCTTTATTCAAAGATGGAACAATAGGAGGAACAATCAACATTGCAATTGTA 1118
Qy 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis 360
Db 1119 GGTCTGATTTCTTAGAAGATGAACAGCCAGGACTGGTGATAGTCACCCAGCAGCCAC 1178
Qy 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 1179 ACCTTAAGTAGCTTCCAGCTGGCAGCTGATGATGGGGAAGATGGGACTGCTCAT 1238
Qy 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db 1239 GACCACGCCACTTACTAGCTGGTCTGGATATATGTTCTCTGGAAGATGAGCCCTGTGAC 1298
Qy 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
Db 1299 ACTTTGGGATTTGCAACCCATAGTGGAAATGTGTAGTAAATATCGCAGCTGCACGATTAA 1358
Qy 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 1359 GAAGATACAGGCTCTGGAGCTGGCCCTTACCATTGCCATGAGCTCGGACACAACTTTGGC 1418
Qy 441 MetIleHisAspGlyGluGlyAsnMetCysLysSerGluGlyAsnIleMetSerPro 460
Db 1419 ATGATTCATGATGGAGAGGGAACATGTGTAAAGTCCGAGGGCAACATCATGTCCTCT 1478
Qy 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
Db 1479 ACATTTGGCAGGACGAATGAGTCTTCTCTGTCACCTCGAGCGCCAGTATCTACAC 1538
Qy 481 LysPheLeuArgSer-ValLysMetProAlaLeu 491
Db 1539 AAATTTCTAAGCACCCGCTCAAGCTATCTGCCTTG 1572

RESULT 7
AR534557 Locus 4042 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 5 from patent US 6734007.
ACCESSION AR534557
VERSION AR534557.1 GI:53924882
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4042)
AUTHORS Fridde,C.J. and Hilbun,E.
TITLE Human proteases and polynucleotides encoding the same
JOURNAL Patent: US 6734007-A 5 11-MAY-2004;
FEATURES
    source          1..4042
                    /organism="unknown"
                    /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 5,71e-147 Length: 4042
Score: 2592.00 Matches: 486
Percent Similarity: 99.33% Conservative: 3
Best Local Similarity: 98.78% Mismatches: 2
Query Match: 98.63% Indels: 1
DB: 6 Gaps: 0

US-10-804-457-2 (1-491) x AR534557 (1-4042)
Qy 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaLeuTrpMetLeuAlaGln 20
Db 99 ATGAAGCCCCCGCGCGGATGGCGGCTTGGCGGCTGTGGATGCTGCTGCGGCGAG 158
Qy 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
```

```
Db 159 GTGGCCGAGCAGGACCTGCGTGGCCATGGACCCGAGCGGCGAGCGCTTGGAGCCCG 218
Qy 41 SerValProArgProProProAlaGluArgProGlyTrpMetGluLysGlyGluTyr 60
Db 219 AGCGTCCCGCGTCTCTCCACCCGCGAGCGCGCGGCTGGATGGAAAAAGGCGCAATAT 278
Qy 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
Db 279 GACCTGGTCTCTGCGCTACAGGTTGACCAACAGGGCGGATACGCTGCTCCCATGAATCATG 338
Qy 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
Db 339 CACCATCAGCGCGGAGAGAGAGAGTGGCGGTCTCCGAGGTGAGTCTCTTCCACTTCGG 398
Qy 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerLeuValAla 120
Db 399 CTGAAGAGCTCCAGCAGCACTTCCACGTGGATCTGAGGACTTCCAGCAGCCCTAGTGGCT 458
Qy 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 459 CTTGGCTTTATTGTGACAGCGTTGGGAAAGACAGGCACTAAGTCTGTGCAGACTTTACCG 518
Qy 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db 519 CCAGAGGACTTCTGTTCTATCAAGGCTTTTGGCATCACACAGAAACTCTCTCAGTGGCC 578
Qy 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluAlaAspTyrPheLeu 180
Db 579 CTTTCACTCCAGCAAGGCTTGTGAGCATGATACGACAGAGAGGAGCAGATTACTTCTTA 638
Qy 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
Db 639 AGGCCACTTCTTCACTGAAACTCTGCGCAGAGCTGCCCAAGGAGCTCGCCA 698
Qy 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
Db 699 TCCACAGTACTGTACAAGAGATCCAAGAGCCCCATGCTCTCTGGGGCCAGTCAGGTCCTG 758
Qy 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db 759 GTGACCTCAGGACATGGAGCTGGCACAATCAACCTTGCACAGCAGGACCTTGCCTG 818
Qy 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
Db 819 GGACTGCCACAAAAGCAGCACTTCTGTGGAAGACGCAAGAAATACATGCCCCAGCTCCC 878
Qy 261 LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db 879 AAGGAAGACCTCTTTCATCTTCCAGATGAGTATAAGTCTTGTCTTACGGCATTAAGCGCTCT 938
Qy 281 LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValAspLys 300
Db 939 CTTCTGAGGTCCCATAGAAATGAAGACTGAACGTGAGGAGACCTTGGTGGTGTGCACAAA 998
Qy 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrThrTyrValLeuThrIleLeuAsn 320
Db 999 AAGATGATGCAAAACCATGGCCATGAAATAATCACCACTAGCTGCTCAGCATCTCAAC 1058
Qy 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAlaIleVal 340
Db 1059 ATGGTATCTGCTTTATTCAAAGATGGAACAATAGGAGGAACAATCAACATTGCAATTGTA 1118
Qy 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis 360
Db 1119 GGTCTGATTTCTTAGAAGATGAACAGCCAGGACTGGTGATAGTCACCCAGCAGCCAC 1178
Qy 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 1179 ACCTTAAGTAGCTTCCAGCTGGCAGCTGATGATGGGGAAGATGGGACTGCTCAT 1238
Qy 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
```



```

Db      1239 GACCAGCCATCTTACTGCTGCTGGATATATGTTCTCGAAGATGAGCCCTGTGAC 1298
Qy      401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAen 420
Db      1299 ACTTTGGGATTTGCAACCCATAGTGGAAATGTGTAGTAAATATCGCAGCTGCACGATTAAT 1358
Qy      421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPhGly 440
Db      1359 GAAGATACAGGCTTGGACTGGCTTACCATTTGCCATGAGTCTGGACACAACTTTGGC 1418
Qy      441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
Db      1419 ATGATTCATGATGGAGAGGACATGTGTAAAGATCCGAGGGCAACATCATGTCCCT 1478
Qy      461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
Db      1479 ACATTTGGCAGGACGCAATGGAGTCTTCTCTGGTCACTCCCTGCAGCCGCGCATATATACAC 1538
Qy      481 LysPheLeuArgSer-VallysMetProAlaLeu 491
Db      1539 AAATTTCTAAGCACCGCTCAAGCTATCTGCCTTG 1572

RESULT 8
CQ850584
LOCUS   CQ850584
DEFINITION Sequence 1053 from Patent EPL447413.
ACCESSION CQ850584
VERSION   CQ850584.1
KEYWORDS  GI:51508796
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS   Tsogai, T., Yamamoto, J., Nishikawa, T., Isono, Y., Sugiyama, T.,
          Otsuki, T., Wakamatsu, A., Ishii, S., Nagai, K. and Irie, R.
TITLE     Full-length human cDNA
JOURNAL   Patent: EP 1447413-A 1053 18-AUG-2004;
          Research Association for Biotechnology (JP)
FEATURES
    source
        1..3287
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.:      8, 92e-147      Length:      3287
Score:          2587.00        Matches:      485
Percent Similarity: 99.19%      Conservative: 3
Best Local Similarity: 98.58%      Mismatches: 3
Query Match:      98.44%      Indels:      1
DB:              6              Gaps:        0

US-10-804-457-2 (1-491) x CQ850584 (1-3287)

Qy      1 MetLysProArgAlaArgGlyTyrArgGlyLeuAlaLeuTrpMetLeuLeuAlaGln 20
Db      139 ATGAAGCCCGCGCGCGATGCGGGCTTGGCGCGCTGTGGATGTCTGTGGCGAG 198
Qy      21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
Db      199 GTGGCCGAGCAGGACCTGTGCGCCATGGGACCCGCGCGCGCGCGCGCGCGCGCG 258
Qy      41 SerValProArgProProProProAlaGluArgProGlyTyrMetGluLysGlyGluTyr 60
Db      259 AGCGTCCCGCGTCTCTCCACCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 318
Qy      61 AspLeuValSerAlaTyrGluValAlaPheHisArgGlyAspTyrValSerHisGluIleMet 80
Db      319 GACCTGGTCTCTGCTACGAGGTTGACACAGGGCGATTACGTCTCCCATGAATCATG 378
Qy      81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100

```

```

Db      379 CACCATCAGCGCGGAGAGAGCAGTGCCTGTCGAGGTTGAGTCTCTTCCACCTTCGG 438
Qy      101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAla 120
Db      439 CTGAAAGGCTCCAGCGACGACCTTCCACGTGGATCTGAGGACTTCCAGCAGCTAGTGCT 498
Qy      121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db      499 CCTGCTTTATTTGTGCAGACGTTGGGAAAGACAGGCCTAAGTCTGTGCAGACTTTACCG 558
Qy      141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db      559 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGCATCACACAGAACTCTCTAGTGCC 618
Qy      161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu 180
Db      619 CTTTCAACCTGCCAAGCTTGTGAGGCATGATACGACACAGAGGCGAGTACTTCTCTA 678
Qy      181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
Db      679 AGGCCACTTCTTTCACACCTCTCATGAAACTCGGAGAGCTGCCAAGCAGCTCGCCA 738
Qy      201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyValAspGluValLeu 220
Db      739 TCCACGCTACTGTACAAGAGATCCACAGAGCCCATGCTCTCTGGGCCAGTGAAGTCTG 798
Qy      221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db      799 GTGACCTCAAGGACATGGGAGCTGGCAGATCAACCTTCACAGCAGCAGCAGCTTCGCTG 858
Qy      241 GlyLeuProGlnLysGlnHisPheCysGlyArgLysLysLysLysLysLysLysLysLys 260
Db      859 GGACTGCCCAAAAGCAGCATTTCTGTGGAAGACCCAGAAATATCATGCCCGGCTGCC 918
Qy      261 LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db      919 AAGGAAGACCTCTTCATCTTGCAGATGAGTATAGTCTTGTACGCGCATGAAGCGCTCT 978
Qy      281 LeuLeuArgSerHisArgAsnGluGluLeuValGluThrLeuValValValAspLys 300
Db      979 CTTCTGAGGTCCATAGAAATGAAGAACTGAACGTGGAGACCTTGGTGGTGGTGGACAAA 1038
Qy      301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrTyrValLeuThrIleLeuAen 320
Db      1039 AAGATGATGCAAAACCATGGCCCATGAAATATATCACCCTACGTGCTCAGCATCTCAAC 1098
Qy      321 MetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAlaIleVal 340
Db      1099 ATGGTATCTGCTTTATTCAAAGATGGAACATAGAGGAAACATCAACATTTGCAATTGTA 1158
Qy      341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisHisAlaAspHis 360
Db      1159 GGTCTGATTTCTTAGAAGATGAACAGCCAGGGCTGGTATAGTCAACACGAGACCCAC 1218
Qy      361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db      1219 ACCTTAAGTAGTCTTCTGCCAGTGGCAGTCTGGATTATGTTCTCGAAGAAATGAGCCCTGTGAC 1278
Qy      381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db      1279 GACCAACCCATCTTACTGACTGGTCTGGATATATGTTCTCGAAGAAATGAGCCCTGTGAC 1338
Qy      401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAen 420
Db      1339 ACTTTGGGATTTGCAACCCATAGTGGAAATGTGTAGTAAATATCGCAGCTGCACGATTAAT 1398
Qy      421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPhGly 440
Db      1399 GAAGATACAGGCTTGGACTGGCTTCCATTTGCCATGAGTCTGGACACAACTTTGGC 1458
Qy      441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460

```

```

Db      1459 ATGATTTCATGAGAGAGGGAACATGTGTAAAAAGTCGAGGGCAACATCATGTGCCCT 1518
Qy      461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTrpLeuHis 480
Db      1519 ACATTGGCAGGACGCAATGGAGTCTTCTCTGGTCACCTGCAGCGGCAGTATCTACAC 1578
Qy      481 LysPheLeuArgSer-ValLysMetProAlaLeu 491
Db      1579 AAATTTCTAAGCACCGCTCAAGCTATCTGCCTTG 1612

RESULT 9
AK127746
LOCUS   Homo sapiens cDNA FLJ45946 fis, clone OCBBF2007039, weakly similar
        to ADAM-TS 7 precursor (EC 3.4.24.-)
ACCESSION
VERSION AK127746
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE  Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T.,
        Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
        Kamiyama,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
        Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Ota,T., Wakamatsu,A.,
        Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K.,
        Nishikawa,T., Kimura,K., Yamaashita,H., Matsuo,K., Nakamura,Y.,
        Sekine,M., Kikuchi,H., Murakawa,K., Kanehori,K.,
        Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
        Sugano,S., Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.
        NEDO human cDNA sequencing project
TITLE   Unpublished
JOURNAL
REFERENCE
AUTHORS Isogai,T. and Yamamoto,J.
TITLE   Direct Submission
JOURNAL
COMMENT
        Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
        Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
        (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
        NEDO human cDNA sequencing project supported by Ministry of
        Economy, Trade and Industry of Japan; cDNA full insert sequencing:
        Research Association for Biotechnology (RAB); cDNA library
        construction: Helix Research Institute (HRI) (supported by Japan
        Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
        HRI, and Biotechnology Center, National Institute of Technology and
        Evaluation; clone selection for full insert sequencing: HRI and
        RAB; annotation: Reverse Proteomics Research Institute, HRI and
        RAB.

FEATURES
        source
        Location/Qualifiers
            1..3287
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="OCBFB2007039"
               /tissue_type="brain"
               /clone_lib="OCBFP2"
               /dev_stage="fetal"
               /note="Cloning vector: pME18SFL3"

ORIGIN
Alignment Scores:
Pred. No.:      8,92e-147      Length:      3287
Score:          2587.00      Matches:      485
Percent Similarity: 99.13%      Conservative: 3
Best Local Similarity: 98.58%      Mismatches: 3
Query Match:    98.44%      Indels:      1
DB:              9      Gaps:          0

US-10-804-457-2 (1-491) x AK127746 (1-3287)

Qy      1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaLeuTrpMetLeuLeuAlaGln 20

```

```

Db      139 ATGAAGCCCCGGCGCGGATGGCGGGCTTGGCGGGCTGTGTGATGCTGTGGCGAG 198
Qy      21 VallalaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaAlaProGlySerPro 40
Db      199 GTGGCCGAGCAGGACCTGCGTGGCCATGGACCCGAGCGGACGCGCTGGAGCCCG 258
Qy      41 SerValProArgProProProProAlaGluArgProGlyTrpMetGluLysGlyGluTrp 60
Db      259 AGCGTCCCGCGTCTCTCCACCCCGGAGCGCGCGGTGGATGGAAAGAGCGCAATAT 318
Qy      61 AspLeuValSerAlaTrpGluValAspHisArgGlyAspTrpValSerHisGluLeuMet 80
Db      319 GACCTGTCTCTGCTACTAGAGTTGACCCACAGGGCGGATTAGCTGTCCCATGAATCATG 378
Qy      81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
Db      379 CACCATCAGCGCGGAGAGAGAGCAGTGCCTGTCCGAGGTTGAGTCTCTTTCACCTTCGG 438
Qy      101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerLeuValAla 120
Db      439 CTGAAAGGCTCCAGCACGACTTCCACGTGGATCTTGAGGACTTCCAGCAGCCTAGTGGCT 498
Qy      121 ProGlyPheLeuValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db      499 CTGGCTTTATTGTGCAGACGTTGGGAAAGACAGCAGCACTAAGTCTGTGCAGACTTTACCG 558
Qy      141 ProGluAspPheCysPheTyrglnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db      559 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGCATCACACAGAAACTCTCTCAGTGGCC 618
Qy      161 LeuSerThrCysGlnGlyLeuSerGlyMetLeuArgThrGluGluAlaAspTrpPheLeu 180
Db      619 CTTTCAACTGCCAAGGCTTGTGAGGCATGTATACGAAACAGAGAGGCGCAGATTACTTCTTA 678
Qy      181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
Db      679 AGGCCACTTCTTTCACACTCTCATGGAAACTCGGCAGAGCTGCCAAGGCGAGCTGCCCA 738
Qy      201 SerHisValLeuTrpLysArgSerThrGluProHisAlaProGlyValaserGluValLeu 220
Db      739 TCCCAAGTACTGTACAGAGATCCACAGAGCCCCATGCTCTCTGGGGCCAGTGCAGTCTCTG 798
Qy      221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db      799 GTGACCTCAAGGACATGGGAGCTGGCACAATCAACCCCTGCACAGCAGGACCTTCGCTCG 858
Qy      241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrcMetProGlnProPro 260
Db      859 GGACTGCCCAAAAGCAGCATTTCTGTGGAAGACCAAGAAATATACATGCCCCAGCTCCC 918
Qy      261 LysGluAspLeuPheLeuProAspGluTrpLysSerCysLeuArgHisLysArgSer 280
Db      919 AAGGAAGACCTTTCATCTTGCAGATCAGTATAGTCTTGTCTTACGCGCATAGCGCTCT 978
Qy      281 LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValValAspLys 300
Db      979 CTTCTAGGTCCTCCATAGAAATGAAGACTGAACGCTGAGACCTTGGTGGTGCACAA 1038
Qy      301 LysMetMetGlnAsnHisGlyHisGluAsnLeuThrThrThrValLeuThrLeuLeuAsn 320
Db      1039 AAGATGATGCAAAACCATGGCCATGAAATATACACCACTAGCTGTCTACGATATCTCAAC 1098
Qy      321 MetValSerAlaLeuPheLysAspGlyThrileGlyGlyAsnileAsnileAlaileVal 340
Db      1099 ATGGTATCTGCTTTATTCAAGATGGAACATAGAGGGAACATCAACATTGCATTGTA 1158
Qy      341 GlyLeuileLeuLeuGluAspGlnProGlyLeuValilleSerHisHisAlaAspHis 360
Db      1159 GGTCTGATTTCTTAGAAGATGAACAGCAGCGGCTGGTGAATAAGTCAACCCAGACCAAC 1218
Qy      361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db      1219 ACCTTAAGTAGTCTTCTGCCAGTGGCAGTCTGGATTGATGGGGAAGATGGGACTCGTCAT 1278

```

QY 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAenGluProCysAasp 400
 Db 1279 GACCAGCCCATTTACTGACTGGTCTGGATATATGTTCTGGAAGATGAGCCCTGTGAC 1338
 QY 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTrpArgSerCysThrIleAen 420
 Db 1339 ACTTTGGGATTTGCACCCATAGTGAATGTAGTAATATCCAGCTGCACGATTAAAT 1398
 QY 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAenPheGly 440
 Db 1399 GAAGATACAGGTCCTGGACTGGCTTACCATTCACCTGCCAGTCTGCACACACTTTGGC 1458
 QY 441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAenIleMetSerPro 460
 Db 1459 ATGATTCATGATGAGAGGAACATGTGTAAAGTCCGAGGCGACATCATGTCCTCCCT 1518
 QY 461 ThrLeuAlaGlyArgAenGlyValPheSerTrpSerProCysSerArgGlnTrpLeuHis 480
 Db 1519 ACATGGCAGACCAATGAGAGTCTTCTCTGGTCACTTGCAGCCGCGCAGTATCTACAC 1578
 QY 481 LysPheLeuArgSer-VallysMetProAlaLeu 491
 Db 1579 AAATTTCTAAGCACCGCTCAAGCTATCTGCCTTG 1612

RESULT 10

HSA315734 3227 bp mRNA linear PRI 01-MAR-2002
 LOCUS
 DEFINITION Homo sapiens mRNA for metalloprotease disintegrin 16 (ADAMTS16 gene).
 ACCESSION AJ315734
 VERSION AJ315734.1 GI:19171177
 KEYWORDS ADAMTS16 gene; disintegrin; metalloprotease; thrombospondin.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Cal.S., Obaya, A.J., Llamazares, M., Garabaya, C., Quesada, V. and Lopez-Otin, C.
 Cloning, expression analysis, and structural characterization of seven novel human ADAMTSs, a family of metalloproteinases with disintegrin and thrombospondin-1 domains
 Gene 283 (1-2), 49-62 (2002)
 21856482
 11867212
 2 (bases 1 to 3227)
 Cal, S.
 Direct Submission
 Submitted (26-JUN-2001) Cal S., Biochemistry and Molecular Biology, University of Oviedo, Campus del Cristo, Asturias. 33006, SPAIN
 Location/Qualifiers
 1. 3227
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 1. 3227
 /gene="ADAMTS16"
 1. 3219
 /gene="ADAMTS16"
 /codon_start=1
 /product="metalloprotease disintegrin 16 with thrombospondin type 1 motif"
 /protein_id="CAC86015.1"
 /db_xref="GI:19171178"
 /db_xref="GOA:Q8TES7"
 /db_xref="UniProt/TREMBL:Q8TES7"
 /translation="MKPRARGWGLAALWMLIAOVAQAPACMGPAAPAAAPGSPVPR PPPAERPGWKEGYDLVSAVEVDHGDYVSEIHHQRRRAVAVSEVESLHLRLK GRPHDPLMLRTSSSIVA PGPIVQLTKTKSVQTLPPDFCFYQGSLSRHSNSVA LSTCOGLSCWIRTEADYFPLPSPHLSKMGRAAGSSPSHLYIKRSTEPHAPGSA VLVSTYRELAKPLHSSDLRLGLPKQHFCCRKKYMPQPKEDLFLPDEYKSLR HKRSLSRHSRNLAVETLVVVVDKQMQNHGHENITTVYVLITLWVSALFKDGTIGN

FEATURES

source
 gene
 CDS

ORIGIN

Alignment Scores:

Pred. No.: 1,32e-146 Length: 3227
 Score: 2584.00 Matches: 484
 Percent Similarity: 99.19% Conservative: 4
 Best Local Similarity: 98.37% Mismatches: 3
 Query Match: 98.33% Indels: 1
 DB: 9 Gaps: 0

US-10-804-457-2 (1-491) x HSA315734 (1-3227)

QY 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaAlaLeuTrpMetLeuLeuAlaGln 20
 Db 1 ATGAAGCCCGCGCGCGGATGGCGGGCTTGGCGCGCTGTGGATGCTGTGGCGCAG 60
 QY 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
 Db 61 GTGCCGAGCAGCACCTCGTCGCGCATGGACCCGCGAGCGCGCTGGAGCCCG 120
 QY 41 SerValProArgProProProAlaGluArgProGlyTrpMetGluLysGlyGluTyr 60
 Db 121 AGGCTCCCGGCTCTCTCCACCCCGGAGCGCGGCTGGATGGAAAGCGCGAATAT 180
 QY 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
 Db 181 GACCTGGTCTCTGCTTACGAGGTTGACCACAGGGCGGATTAGTGTCCCATGAATCATG 240
 QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
 Db 241 CACCATCAGCGCGAGAGCAGTGGCTGCGAGGTGAGTCTCTTCACTTCGG 300
 QY 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAla 120
 Db 301 CTGAAAGGCCCGCCAGCACGACTTCACATGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360
 QY 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
 Db 361 CTTGGCTTATTTGTGCAGAGCTTGGGAAGACAGGCCTAAGTCTGTGCAGACTTTACCG 420
 QY 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAenSerSerValAla 160
 Db 421 CCAGGAGACTTCTGTTTCTATCAAGGCTCTTTCGATCATCACAGAACTCTCTCAGTGGCC 480
 QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaPheTyrPheLeu 180
 Db 481 CTTTCAACCTGCCAGGCTTGT CAGGCATGATACGAACAGAGGAGCAGATCTACTTCTTA 540
 QY 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
 Db 541 AGGCCACTTCTCTTACACCTCTCATGGAACTCGGCAGAGCTGCCCAAGCAGCTCGCCA 600
 QY 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
 Db 601 TCCACGCTACTGTACAGAGATCCACAGAGCCCCATGTCTCTGGGCCAGTGTGAGTCTCTG 660
 QY 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
 Db 661 GTGACCTCAAGGACATGGGAGCTGGCACATCAACCCCTGCACAGCAGGACCTTTCGCTTG 720

```
QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
Db 721 GGACTGCCACAAAGCAGCATTTCTGTGGAAACCAAGAAATACATGCCCGCCTCCC 780

QY 261 LysGluAspLeuPheLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db 781 AAGGAAGACCTCTTCATCTTGCAGATGATATAGTCTTGTTCAGGATTAAGCGCTCT 840

QY 281 LeuLeuArgSerHisArgAsnGluGluLeuAenValGluThrLeuValValAspLys 300
Db 841 CTTCTGAGGTCCTATAGAAATGAAGAACTGAACGTTGGAGACCTTGTGTGGTGACAAA 900

QY 301 LysMetMetGlnAsnHisGlyHisGluAenIleThrTyrValLeuThrIleLeuAen 320
Db 901 AAGATGATGCAAAACCATGGCCCATGAAATATACACCACTACGCTGCTCAGGATCTCAAC 960

QY 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAlaIleVal 340
Db 961 ATGGTATCTGCTTTATTCAGATGGAACATAGAGGAAACATCAACATTCATTTGTA 1020

QY 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis 360
Db 1021 GGTCTGATTTCTTAGAAGATGAACAGCCAGGACTGTGTGATAAGTCAACCAAGACAC 1080

QY 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 1081 ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCAT 1140

QY 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db 1141 GACCAGCCATCTTACTGACTGGTCTGGATATATGTTCTCTGGAAGATGAGCCCTGTGAC 1200

QY 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAen 420
Db 1201 ACTTTGGATTTGCCACCATAGTGAATGTGTAGTAAATATCGCAGCTGCAGCATTAAT 1260

QY 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 1261 GAAATACAGGTCTTTGGACTGGCTTCCACATGTCCTGATGCTGAGTCTGGACACAACTTTGGC 1320

QY 441 MetIleHisAspGlyGluClyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
Db 1321 ATGATTCATAGTGGAGAGGGAACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCT 1380

QY 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
Db 1381 ACATTGGCAGGAGCAATGGAGTCTTCTCTGTCACCTGCGAGCGCCAGCATCTACAC 1440

QY 481 LysPheLeuArgSer-ValLysMetProAlaLeu 491
Db 1441 AAATTTCTAAGCAGCGCTCAAGCTATCTGCCTTG 1474

RESULT 11
AX319852 3675 bp DNA linear PAT 14-DEC-2001
LOCUS Sequence 16 from Patent WO0183782.
DEFINITION AX319852
ACCESSION AX319852
VERSION AX319852.1 GI:17901442
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Plowman,G.D., Whyte,D., Sudarsanam,S., Manning,G., Caenepeel,S. and
Payne,V.
TITLE Novel proteases
JOURNAL Patent: WO 0183782-A 16 08-NOV-2001;
Sugen, Inc. (US)
FEATURES
source
1..3675
Location/Qualifiers
/mol_type="unassigned DNA"
```

```
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1,55e-146 Length: 3675
Score: 2584.00 Matches: 484
Percent Similarity: 99.19% Conservative: 4
Best Local Similarity: 98.37% Mismatches: 3
Query Match: 98.33% Indels: 1
DB: 6 Gaps: 0

US-10-804-457-2 (1-491) x AX319852 (1-3675)
QY 1 MetLysProArgAlaArgGlyTyrPargGlyLeuAlaLeuTyrMetLeuLeuAlaGln 20
Db 1 ATGAAGCCCGCGCGCGCGGATGGCGGGCTTGGCGGCGCTGTGGATGCTGTGGCGCAG 60

QY 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaAalaProGlySerPro 40
Db 61 GTGGCCGAGCAGGACCTGCGTGCCTCATGGACCCCGCAGCGCGCTGGAGCCCG 120

QY 41 SerValProArgProProProAlaGluArgProGlyTyrMetGluLysGlyGluTyr 60
Db 121 AGCGTCCCGCTCTCTCCACCGCGGCGCGCGCTGGATGMAAAGGCGCAATAT 180

QY 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluLeuMet 80
Db 181 GACCTGGTCTCTGCTAGCAGGTTGACCACAGCGGCGGATACGTCTCCCATGAAATCATG 240

QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuAeq 100
Db 241 CACCATCAGCGCGGAGAGAGCAGTGGCGGTGTCGAGGTTGAGTCTCTTCCACTTCGG 300

QY 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerLeuValAla 120
Db 301 CTGAAGGCCCGCCAGCAGCTTCCACATGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360

QY 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 361 CTTGGCTTTATTTGTGCAGACGTTGGGAAAGACAGGCACCTAAGTCTGTGCAGACTTTACCG 420

QY 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db 421 CCAGAGGACTTCTGTCTTATCAAGGCTTTTGGCATGATACGAAACAGAGAGGAGGAGG 480

QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu 180
Db 481 CTTTCNACCTGCCAAGGCTTGTGAGGATGATACGAAACAGAGAGGAGGAGGAGGAGG 540

QY 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
Db 541 AGGCCACTTCTTCCACACTCTCATGGAAACTGGGAGAGCTGCCCAAGGAGCTCGCCA 600

QY 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyLysSerGluValLeu 220
Db 601 TCCCAAGTACTGTACAGAGATGCCAGAGCCCCATGCTCTCTGGGCGCAGTGGAGTCTCG 660

QY 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db 661 GTGACCTCAGGACATGGAGCTGGCAGATCAACCCCTGCACAGGAGGAGGAGGAGGAGG 720

QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
Db 721 GGACTGCCACAAAGCAGCATTTCTGTGGAAAGCAAGAAATACATGCCCCAGGCTCC 780

QY 261 LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db 781 AAGGAAGACCTTCTTCATCTTCCAGATGATATAGTCTTGTCTTACGCGATCAAGCGCTCT 840

QY 281 LeuLeuArgSerHisArgAsnGluGluLeuAenValGluThrLeuValValAspLys 300
Db 841 CTTCTGAGGTCCTATAGAAATGAAGAACTGAACGTTGGAGACCTTGTGTGGTGACAAA 900
```

```

QY 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrThrTyrValLeuThrIleLeuAsn 320
Db 901 AAGATGATGCAAAACCATGGCCATGAAATATACCACTACGTCGTGCAGTACTCAAC 960
QY 321 MetValSerAlaLeuPheLeuAspGlyThrIleGlyGlyAsnIleAlaIleVal 340
Db 961 ATGGTATCTGCTTTATTCAAGATGGAAACATAGGAGGAACATCAACATTGCAATTGTA 1020
QY 341 GlyLeuIleLeuLeuGluAspGlnProGlyLeuValIleSerHisHisAlaAspHis 360
Db 1021 GGTCGTGATCTCTTAGAAGATGAACACAGGAGCTGGTGATAAGTACACGAGAGCCAC 1080
QY 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 1081 ACCTTAAGTAGTCTCTGCCAGTGGCAGTCTGGATTGATGGGAAGATGGGACTCTGCAIT 1140
QY 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysGlnProCysAsp 400
Db 1141 GACCACGCCATCTTACTGACTGGTCTGGATATATGTTCTGGGAAGATGAGCCCTGTGAC 1200
QY 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
Db 1201 ACTTTGGGATTTCACCCATAAGTGGAAATGTGTAGTAATATCGCAGCTGCAGATTAT 1260
QY 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 1261 GAAGATACAGGCTCTGGACTGGCTTTACCAATGGCCATGCTGGACACAACTTTGGC 1320
QY 441 MetIleHisAspGlyGluGlyAsnMetCysValLeuSerGluGlyAsnIleMetSerPro 460
Db 1321 ATGATTCATGTGAGAGGAAACATGTGTAAAGTCCGAGGGCAACATCATGTCCCT 1380
QY 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
Db 1381 ACATTGGCAGGACGCAATGAGTCTTCTCTGCTCACCTCGCAGCGCCAGTATCTACAC 1440
QY 481 LysPheLeuArgSer-ValLysMetProAlaLeu 491
Db 1441 AATTTCTAAGCACCGCTCAGCTATCTGCTTG 1474

RESULT 12
BD170080
LOCUS
DEFINITION Novel ADAMTS family polypeptide and gene encoding the same.
ACCESSION BD170080
VERSION BD170080.1
KEYWORDS WO 0231163-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ohara O., Nagase, T., Nomura, N., Yano, K., Murakami, K., Yasuda, S. and
Kanzaki, K.
Novel ADAMTS family polypeptide and gene encoding the same
Patent: WO 0231163-A 2 18-APR-2002;
KAZUSA DNA RESEARCH INSTITUTE, MITSUBISHI PHARMA CORP, OSAMU OHARA,
TAKAHIRO NAGASE, NOBUO NOMURA, KAZUHIRO YANO, KOJI MURAKAMI, SHINICHIRO
YASUDA, KOJI KANZAKI
OS Homo sapiens (human)
PN WO 0231163-A/2
PD 18-APR-2002
PF 11-OCT-2001 WO 2001JP008913
PR 11-OCT-2000 JP OOP 311309, 02-APR-2001 JP OIP 102905 PI
OSAMU OHARA, TAKAHIRO NAGASE, NOBUO NOMURA, KAZUHIRO YANO, KOJI PI
MURAKAMI
PI SHINICHIRO YASUDA, KOJI KANZAKI
PC C12N15/57, C12N9/64, C12N5/00, C12N1/19, C12N1/21, C12N1/15, C07K16/
PC 40, C12Q1/68,
PC G01N33/50
CC Novel ADAMTS family polypeptide and gene encoding the same FH
Key Location/Qualifiers
FT CDS (770)..(4444).

```

```

FEATURES
Source
Location/Qualifiers
1..5610
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 2,58e-146 Length: 5610
Score: 2584.00 Matches: 484
Percent Similarity: 99.19% Conservative: 4
Best Local Similarity: 98.37% Mismatches: 3
Query Match: 96.33% Indels: 1
DB: 6 Gaps: 0

US-10-804-457-2 (1-491) x BD170080 (1-5610)

QY 1 MetLysProArgAlaArgGlyTyrArgGlyLeuAlaLeuTyrMetLeuLeuAlaGln 20
Db 770 ATGAAGCCCCCGCGCGCGGATGGCGGGCTTGGCGGCGCTGTGGATGCTGCTGGCGCAG 829
QY 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaAlaProGlySerPro 40
Db 830 GTGGCCGAGCAGGACCTGGGTGGCCATGGGACCCGCGGCGGCGCTGGGAGGCGCG 889
QY 41 SerValProArgProProProProAlaGluArgProGlyTyrMetGluLysGlyGluTyr 60
Db 890 AGCGTCCCGGCTCTCTCCACCCGCGGAGCGCGGCTGGATGGAAAAGGCGCATAT 949
QY 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
Db 950 GACCTGCTCTCGCTACGAGGTGACCAAGGGGCGGATACGTGTCCCATGAATCATG 1009
QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
Db 1010 CACCATCAGCGCGGAGAGAGCAGTGGCGGTGTCGAGGTTGAGTCTCTTACCTTCG 1069
QY 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAla 120
Db 1070 CTGAAGGCCCGCCAGGACGACATTCACATGGATCTCGAGGACTTCCAGGACCTAGTGGCT 1129
QY 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 1130 CCTGGCTTTATTGTGCAGAGCTTGGGAAAGACAGGACCTAAGTCTGTGCAGACTTTACC 1189
QY 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db 1190 CCAGAGGACTTCTGTTCTTATCAAGGCTCTTTGCGATCATCACAGAAACTCTCTCAGTGGCC 1249
QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu 180
Db 1250 CTTTCAACCTGCGCAAGGCTTGTGAGCATGATACGAACAGAAGAGCGAGATTACTCTCTA 1309
QY 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
Db 1310 AGGCCACTTCTTCTCACCTCTCATGTGAAACTCGCGCAGAGCTGCCCAAGGAGCTCGCCA 1369
QY 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
Db 1370 TCCGACGTACTGTACAGAGATCCACAGGCCCCATGCTCTCTGGGGCGGAGGAGTCTCTG 1429
QY 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db 1430 GTGACCTCAAGGACATGGGAGCTGGACATCAACCCCTGCACAGCAGCAGCCTTCGCTG 1489
QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
Db 1490 GGACTGCCACAAAAGCAGCATTTCTGTGGAAGAGCGCAAGAAATACATGCCCGAGCTCCC 1549
QY 261 LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisArgSer 280
Db 1550 AAGGAAGACCTTCTTCACTCTTCCAGATGAGTATTAAGTCTTGTACGCGCATAGCGCTCT 1609

```

Qy	281	LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValValAspLys	300
Db	1610	CTTCTGAGGTCCCATAGAAATGAAGAACTGAACGCTGGAGACCTTTGGTGGTGCACAAA	1669
Qy	301	LysMetMetGlnAsnHisGlyHisGluAsnIleThrThrTyrValLeuThrIleLeuAsn	320
Db	1670	AAGATGATGCAAAACCATGGCCATGAATAATCACCACCTAGTGTCTCAGATACTCAAC	1729
Qy	321	MetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAsnIleAlaIleVal	340
Db	1730	ATGGTATCTGCTTTATTCAAAGATGGAACAATAGGAGGAACATCAACATTGCAATTGTA	1789
Qy	341	GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis	360
Db	1790	GGTCTGATTTCTTAGAAGATGAACAGCCAGGACTGGTGTATAAGTCACCCAGCAGACCAC	1849
Qy	361	ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis	380
Db	1850	ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGGAAGATGGGACTGTCAT	1909
Qy	381	AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp	400
Db	1910	GACCACGCCATCTTACTGACTGGTCTGGATATATGTTCTTGGAGAATGAGCCCTGTGAC	1969
Qy	401	ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn	420
Db	1970	ACTTTGGGATTTGGACCCATTAAGTGAATGTGTAGTAAATATCGCAGCTGCACGATTAA	2029
Qy	421	GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly	440
Db	2030	GAAGATACAGGTCTTGGACTGGCCCTTCCACCATTGGCCATGAGTCTGGACACAACCTTTGGC	2089
Qy	441	MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro	460
Db	2090	ATGATTTCATGTGGAGAGGGAACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCT	2149
Qy	461	ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis	480
Db	2150	ACATTTGGCAGCAGCAATGGAGTCTTCTCCTGGTCACTTGCAGCGCCAGTATCTACAC	2209
Qy	481	LysPheLeuArgSerValLysMetProAlaLeu	491
Db	2210	AAATTTCTAAGCACCGCTCAAGCTATCTGCCTTG	2243
RESULT 13			
LOCUS	BD170083	8435 bp	DNA linear PAT 17-JAN-2003
DEFINITION	Novel ADAMTS family polypeptide and gene encoding the same.		
ACCESSION	BD170083		
VERSION	BD170083.1		
KEYWORDS	GI:27875895		
SOURCE	WO 0231163-A/5.		
ORGANISM	synthetic construct		
REFERENCE	synthetic construct		
AUTHORS	other sequences; 1 (bases 1 to 8435)		
TITLE	Chara.O., Nagase.T., Nomura.N., Yano.K., Murakami.K., Yasuda.S. and		
JOURNAL	Kanzaki.K.		
COMMENT	Novel ADAMTS family polypeptide and gene encoding the same		
	Patent: WO 0231163-A 5 18-APR-2002;		
	KAZUSA DNA RESEARCH INSTITUTE,MITSUBISHI PHARMA CORP,OSAMU OHARA,		
	TAKAHIRO NAGASE,NOBUO NOMURA,KAZUHIRO YANO,KOJI MURAKAMI,SHINICHIRO		
	YASUDA, KOJI KANZAKI		
	OS Artificial Sequence		
	PN WO 0231163-A/5		
	PD 18-APR-2002		
	PF 11-OCT-2001 WO 2001JP008913		
	PR 11-OCT-2000 JP 00P 311309,02-APR-2001 JP 01P 102905 PI		
	OSAMU OHARA, TAKAHIRO NAGASE, NOBUO NOMURA, KAZUHIRO YANO, KOJI PI		
	MURAKAMI		
	PI SHINICHIRO YASUDA, KOJI KANZAKI		
	PC C12N15/57,C12N9/64,C12N5/00,C12N1/19,C12N1/21,C12N1/15,C07K16/		
	PC 40,C12Q1/68,		
	PC G01N33/50		


```
Db 4772 GGACTGCCACAAAGCAGCATTTCTGTGGAGACCGCAAGAAATACATGCCCGCCTCCC 4831
Qy 261 LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db 4832 AAGGAAGACCTTTATCTTGCAGATAGATATAAGTCTTGTACGGCATAAAGCGCTCT 4891
Qy 281 LeuLeuArgSerHisArgAsnGluLeuLeuValGluThrLeuValValValAspLys 300
Db 4892 CTTCTGAGGTCCCATAGAAATGAAAGACTGAACGTGGAGACTTGGTGGTGGCAGAA 4951
Qy 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrThrTyrValLeuThrIleLeuAsn 320
Db 4952 AAGATGATGCAAAACCATGGCCATGAAATATACACCTACGTCTCAGATACACTCAAC 5011
Qy 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyValAsnIleAsnIleAlaIleVal 340
Db 5012 ATGGTATCTGCTTTATTCAAGATGGAAACATAGGAGGAACATCAACATGTGCAATTGTA 5071
Qy 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis 360
Db 5072 GGTCTGATTTCTTAGAAGATGAACACCGCAGACTGGTGATAAGTCACCGCAGACCAAC 5131
Qy 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 5132 ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCAT 5191
Qy 381 AsnHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db 5192 GACCACCCCATCTTACTGACTGTCTGGATATATATGTTCTTGGAGAAATGAGCCCTGTGAC 5251
Qy 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
Db 5252 ACTTTGGATTTGACCCCATAGTGGAAATGTGTAGTAAATATCGCAGCTGCACGATTAA 5311
Qy 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 5312 GAAGATACAGGCTTGGACTGGCTTACCATGTCCTGATGAGTCTGGACACACTTTGGC 5371
Qy 441 MetIleHisAspGlyGluGlnMetCysValLysLysSerGluGlyAsnIleMetSerPro 460
Db 5372 ATGATTCATGATGGAGAGGAAACATGTGTAAAGTCCGAGGGCAACATCATGTGCCCT 5431
Qy 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
Db 5432 ACATTGGCAGGACGAATGGAGTCTTCTCTGGTCCACCTGCACGCCCGCAGTATCTACAC 5491
Qy 481 LysPheLeuArgSerValLysMetProAlaLeu 491
Db 5492 AAATTTCTAAGCACCGCTCAAGCTATCTGCCTTG 5525

RESULT 14
BD170084
LOCUS BD170084
DEFINITION Novel ADAMTS family polypeptide and gene encoding the same.
ACCESSION BD170084
VERSION BD170084.1 GI:27875896
KEYWORDS WO 0231163-A/6.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 8505)
AUTHORS Ohara,O., Nagase,T., Nomura,N., Yano,K., Murakami,K., Yasuda,S. and Kanaki,K.
TITLE Novel ADAMTS family polypeptide and gene encoding the same
JOURNAL Patent: WO 0231163-A 6 18-APR-2002;
KAZUSA DNA RESEARCH INSTITUTE,MITSUBISHI PHARMA CORP OSAMU OHARA,
TAKAHIRO NAGASE,NOBUO NOMURA,KAZUHIRO YANO,KOJI MURAKAMI,SHINICHIRO
YASUDA,KOJI KANZAKI
OS Artificial Sequence
PN WO 0231163-A/6
PD 18-APR-2002
PF 11-OCT-2001 WO 2001JP008913
PR 11-OCT-2000 JP ODP 311309,02-APR-2001 JP OIP 102905 PI
```

```
OSAMU OHARA, TAKAHIRO NAGASE, NOBUO NOMURA, KAZUHIRO YANO, KOJI PI
MURAKAMI, SHINICHIRO YASUDA, KOJI KANZAKI
PC C12N15/57, C12N19/64, C12N5/00, C12N1/19, C12N1/21, C12N1/15, C07K16/
PC 40, C12Q1/68,
PC G01N33/50
CC Description of Artificial Sequence: base sequence of the CC
Plasmid
CC pFastBac1-MS/HT-PJ01256-2
FH Key Location/Qualifiers
FT source 1..8505
FT /organism='Artificial Sequence'.
FT /location=Qualifiers
FT /organism='synthetic construct'
FT /mol_type='genomic DNA'
FT /db_xref='taxon:32630'

FEATURES
source
Pred. No.: 4.25e-146 Length: 8505
Score: 2584.00 Matches: 484
Percent Similarity: 99.19% Conservative: 4
Best Local Similarity: 98.37% Mismatches: 3
Query Match: 98.33% Indels: 1
DB: Gaps: 0
US-10-804-457-2 (1-491) x BD170084 (1-8505)

ORIGIN
Alignment Scores:
Pred. No.: 4.25e-146 Length: 8505
Score: 2584.00 Matches: 484
Percent Similarity: 99.19% Conservative: 4
Best Local Similarity: 98.37% Mismatches: 3
Query Match: 98.33% Indels: 1
DB: Gaps: 0

Qy 1 MetLysProArgAlaArgGlyTyrArgGlyLeuAlaLeuTrpMetLeuLeuAlaGln 20
Db 4122 ATGAGCCCCCGCGCGCGGATGGCGGCGCTGTGGCGCTGTGGCGCAG 4181
Qy 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
Db 4182 GTGCGCGAGCAGGCACCTCGGTGGCGCATGGGACCGCGCGCGCGCGCGCGCG 4241
Qy 41 SerValProArgProProProProAlaGluArgProGlyTyrMetGluLysGlyTyr 60
Db 4242 AGCGTCCCCCGCTCTCTCCACCCGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 4301
Qy 61 AspLeuValSerAlaTyrGluValAlaPheHisArgGlyAspTyrValSerHisGluLeuMet 80
Db 4302 GACCTGGTCTCTGCTACGAGGTTGACACAGGGCGGATTACGTGCCAATAATCATG 4361
Qy 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
Db 4362 CACCATCAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4421
Qy 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAla 120
Db 4422 CTGAAAGGCCCGCAGGCACGACTTCCACATGATCTGAGGACTTCCAGCAGCTTAGTGCT 4481
Qy 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 4482 CTTGGCTTTATTGTGACAGCGTTGGGAAAGACAGCGCACTAAGTCTGTGCGACTTTACCG 4541
Qy 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db 4542 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGATCACACAGAAACTCCTCAGTGGCC 4601
Qy 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluAlaAspTyrPheLeu 180
Db 4602 CTTTCAACCTGCCAAGGCTTGTGAGGCATGATACGAACAGAGAGAGAGAGAGAGAGAG 4661
Qy 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaGlnGlySerSerPro 200
Db 4662 AGGCCACTTCTCTTCCACACCTCTCATGGAAACTCGGAGAGCTGCCCAAGCGCAGCTGCCA 4721
Qy 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
Db 4722 TCCACAGTACTGTACAAGAGATCCACAGAGGCCCAATGCTCTCTGGGCGCAGTGGGCTCTG 4781
```



```
QY 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db 4782 GTGACCTCAAGACATGGAGCTGGCACAATCAACCTTCACAGCAGCAGCTTCGCCTG 4841
QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgGlyValLysLysTyrMetProGlnProPro 260
Db 4842 GGACTGGCCACAAAGACGATTTCTGTGGAAAGCGCAGAAATACATGCCCCAGCTCC 4901
QY 261 LysGluAspLeuPheLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db 4902 AAGGAAGACCTCTTCATCTTCAGATGAGTATAGTCTTGTCTACGGCATAGCGCTCT 4961
QY 281 LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValAspLys 300
Db 4962 CTTCTGAGGTCCCATAGAAATGAAGAACTGAACGTGGAGACCTTGTGTGTGTGCACAA 5021
QY 301 LysMetMetGlnAsnHisGlyHisGluLeuLeuThrThrTyrValLeuThrLeuLeu 320
Db 5022 AAGATGATGCAAAACCATGGCCATGAAATATACCACTACCTGCTACGATATCTCAAC 5081
QY 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAlaIleVal 340
Db 5082 ATGGTATCTCTTTATCAAGATGGAACAATAGGAGGAACAATCAACATTGCAATTGTA 5141
QY 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisHisAlaAspHis 360
Db 5142 GGTCTGATTTCTTAGAAGATGAACAGCCAGGACTGGTGTATAGTCAACAGCAGACCA 5201
QY 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 5202 ACCTTAAGTAGCTTCTGCCAGTGGCAGCTCTGGATTGATGGGGAAGATGGGACTCGTCA 5261
QY 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db 5262 GACCACGCCATCTTACTGCTGCTGATATATGTTCTCGAAGATGAGCCCTGTGAC 5321
QY 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
Db 5322 ACTTTGGGATTTGCAACCCATAAGTGAATGTGTAGTAAATATCGCAGCTGCACGATTAA 5381
QY 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 5382 GAAGATACAGGTCTTGGAGCTGGCCCTTCCATGTGCCCATGAGTCTGGACACAACTTTGG 5441
QY 441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
Db 5442 ATGATTATCATGGAGAGGAAACATGTGTAAAGTCCGAGGGCAACATCATGTCCCT 5501
QY 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
Db 5502 ACATTTGGCAGGACGCAATGGAGTCTTCTCTGTCACCTGCGCCGCGCAGTATCTACAC 5561
QY 481 LysPheLeuArgSerValLysMetProAlaLeu 491
Db 5562 AAATTTCTAAGCACCGCTCAAGCTATCTGCCTTG 5595

RESULT 15
LOCUS CQ738549
DEFINITION Sequence 24483 from Patent WO02068579.
ACCESSION CQ738549
VERSION CQ738549.1 GI:42339689
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, B.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 24483 06-SEP-2002;
```

```
PE Corporation (NY) (US)
Location/Qualifiers
1. 1313
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 2,18e-131 Length: 1313
Score: 2324.00 Matches: 435
Percent Similarity: 99.77% Conservative: 1
Best Local Similarity: 99.54% Mismatches: 1
Query Match: 88.43% Indels: 0
DB: 6 Gaps: 0

US-10-804-457-2 (1-491) x CQ738549 (1-1313)

QY 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaLeuTrpMetLeuLeuAlaGln 20
Db 1 ATGAAGCCCGCGCGCGCGGATGGCGGGCTTTGGCGGCGCTGTGGATGCTGCTGGCGCAG 60
QY 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaAlaProGlySerPro 40
Db 61 GTGGCCGAGCAGCAGCAGCTGCGCGCCATGGACCCCGCAGCGCGCGCTGGAGCCCG 120
QY 41 SerValProArgProProProAlaGluArgProGlyTrpMetGluLysGlyGluTyr 60
Db 121 AGCGTCCCGCGTCTCTCCACCCGCGGCGCGCGCGCTGGATGGAAAGGCGCAATAT 180
QY 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
Db 181 GACCTGGTCTCTGCGCTACGAGGTTGACCACAGGGCGGATACGTGTCCCATGAAATCATG 240
QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
Db 241 CACCATCAGCGCGGAGAGAGAGAGAGTGGCGGTCCGAGGTGAGTCTCTTCCACTTCGG 300
QY 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerLeuValAla 120
Db 301 CTGAAGGCCCGCAGCAGCAGCTTCCACATGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360
QY 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 361 CTGGCTTTATTTGTCAGACGTTGGGAAAGACAGGCACCTAAGTCTGTGAGACTTTACCG 420
QY 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db 421 CCAGAGGACTTCTGTGTTCTATCAGGCTCTTTGGCATCACACAGAAACTCTCCAGTGGCC 480
QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluAlaAspTyrPheLeu 180
Db 481 CTTTCNACTGCCAAGGCTTGTACGATGATACGACAGAGAGAGCAGATTTACTTCTTA 540
QY 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
Db 541 AGGCCACTCTCTCACACTCTCTATGAAACTCGCGCAGAGCTGCCCAAGCGCAGCTCGCCA 600
QY 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyLysSerGluValLeu 220
Db 601 TCCACAGTACTGTACAGAGATCCACAGAGCCCCATGCTCTCTGGGGCCAGTGGAGTCTCG 660
QY 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db 661 GTGACTCAAGGACATGGAGCTGGCACAATCAACCTTCACAGCAGCAGCAGCTTCGCGCTG 720
QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
Db 721 GGACTGCCACAAAAGCAGCATTTCTGTGGAAAGCGCAAGAAATACATGCCCCAGCGCTCC 780
QY 261 LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db 781 AAGGAAGACCTCTTCTCATCTTCCAGATGATAGTCTTGTGTCGCGCATAGCGCTCT 840
```

```
Qy 281 LeuLeuArgSerHisArgAsnGluGluLeuAenValGluThrLeuValValValAspLys 300
Db 841 CTTCTGAGGTCCCATAGAAATGAAGAACTGAAACGTGGAGACCTTGGTGGTCGACAAA 900
Qy 301 LysMetMetGlnAsnHisGlyHisGluAenIleThrThrTyValLeuThrIleLeuAen 320
Db 901 AAGATGATGCAAAACCATGGCCATGAATATCACCACCTACGTGCTCAGATACTCAAC 960
Qy 321 MetValSerAlaLeuPheIysAspGlyThrIleGlyGlyAenIleAenIleAlaIleVal 340
Db 961 ATGGTATCTGCTTTTATTCAAGATGGAAACAATAGGAGGAAACATCAACATTGCAATTGTA 1020
Qy 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisHisAlaAspHis 360
Db 1021 GGTCTGATTCTCTAGAGATGAACAGCCAGACTGGTGATAGTCAACACGAGACCAC 1080
Qy 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 1081 ACCTTAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCAT 1140
Qy 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db 1141 GACCACGCCATCTTACTGACTGGTCTGGATATATGTTCTTGGAAAGATGAGCCCTGTGAC 1200
Qy 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyTrpArgSerCysThrIleAen 420
Db 1201 ACTTTGGGATTGACCCATAGTGGAAATGTGTAGTAATATCGCAGCTGCACGATTAAAT 1260
Qy 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHis 437
Db 1261 GAAGATACAGGTCTTGGACTGGCTTCACCATTTGCCCTGAGTCTGGACAC 1311
```

Search completed: June 14, 2005, 13:09:05
Job time : 5268 secs

GenCore-version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 13, 2005, 21:26:38 ; Search time 669 Seconds
(without alignments)
4344.683 Million cell updates/sec

Title: US-10-804-457-2

Perfect score: 2628

Sequence: 1 MKPRARGWRGLAALMLLQA.....SPCSRQYLHLFSLRSVMPAL 491

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n.model -DEV=xlp

-Q=/cgn2_1/USPTO_spool_p/US10804457/runat_13062005_173508_3493/app_query.fasta_1.647
-DB=N Geneseq 16Dec04 -OFT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOFCU=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10804457@cgn 1 1 708 @runat_13062005_173508_3493 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2628	100.0	1476	9	Ada50759 cDNA enco
2	2628	100.0	1476	10	Ab557766 cDNA enco
3	2592	98.6	3675	6	Aal43654 Human MDT
4	2592	98.6	3675	9	Ada50761 cDNA enco
5	2592	98.6	3675	10	Ab557767 Novel hum

6	2592	98.6	3675	12	ACC85689	Acc85689 Human pro
7	2592	98.6	4042	9	ADA50763	Ada50763 Novel hum
8	2592	98.6	4042	10	AB557768	Ab557768 cDNA enco
9	2587	98.4	3287	13	ADR07547	Adr07547 Full leng
10	2584	98.3	3227	10	ACC00343	Acc00343 Human ADA
11	2584	98.3	3675	6	AAS97174	Aas97174 Human met
12	2584	98.3	3675	9	ACF57047	Acf57047 Human ADA
13	2584	98.3	5610	6	ABK49822	Abk49822 Human CDN
14	2584	98.3	8435	6	ABK49825	Abk49825 Plasmid p
15	2584	98.3	8505	6	ABK49826	Abk49826 Plasmid p
16	1934	73.6	2997	6	AB559323	Ab559323 Human zin
17	1934	73.6	2997	10	ADJ38419	Adj38419 Human CDN
18	1929	73.4	2902	6	AB559325	Ab559325 Human ADA
19	1929	73.4	2902	10	ADJ38423	Adj38423 Human CDN
20	1791.5	68.2	2166	8	ADA50516	Ada50516 Human pro
21	1707	65.0	4888	6	ABK86137	Abk86137 cDNA enco
22	1707	65.0	4902	13	ACN43153	Acn43153 Human dia
23	1683	64.0	2895	6	AB559326	Ab559326 Human ADA
24	1683	64.0	2895	10	ADJ38425	Adj38425 Human CDN
25	1511	57.5	4234	6	ABK49821	Abk49821 Human CDN
26	1300	49.5	3630	12	ADP69016	Adp69016 Human NOV
27	1300	49.5	3815	10	ADD22494	Add22494 ADAMTS fa
28	1299	49.4	3630	12	ADP69020	Adp69020 Human NOV
29	1298	49.4	3013	8	ABZ79549	Abz79549 Human pro
30	1298	49.4	3013	9	ACD25896	Acd25896 Novel hum
31	1298	49.4	3389	7	ADR43746	Adr43746 Human pro
32	1298	49.4	3630	8	ABT16078	Abt16078 NOVX rela
33	1298	49.4	3630	12	ADP69018	Adp69018 Human NOV
34	1298	49.4	5576	10	ADD22492	Add22492 ADAMTS fa
35	1297	49.4	2217	8	ABZ79551	Abz79551 Human pro
36	1297	49.4	2217	9	ACD25898	Acd25898 Novel hum
37	1297	49.4	3630	12	ADP69022	Adp69022 Human NOV
38	1295.5	49.3	1986	9	AD57932	Aad57932 Human agg
39	1295.5	49.3	3663	9	AD57932	Aad57932 Human agg
40	1295	49.3	3630	12	ADO41757	Ado41757 Novel hum
41	1293.5	49.2	2589	8	ABZ79548	Abz79548 Human pro
42	1293.5	49.2	2589	9	ACD25895	Acd25895 Novel hum
43	1293.5	49.2	3666	6	ABN89277	Abn89277 Human agg
44	1292.5	49.2	1530	8	ABZ79550	Abz79550 Human pro
45	1292.5	49.2	1530	9	ACD25897	Acd25897 Novel hum

ALIGNMENTS

RESULT 1
ADA50759
ID ADA50759 standard; cDNA; 1476 BP.
XX
AC ADA50759;
DT 20-NOV-2003 (first entry)
XX
DE cDNA encoding novel human protein (NHP) #1.
XX
KW Novel human protein; NHP; protease; zinc metalloproteinase; gene therapy;
XX human; gene; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT CDS 1..1476
FT FT /*tag= a
FT FT /product= "NHP #1"
XX
XX US2002193583-A1.
XX
XX PD 19-DEC-2002.
XX
XX PF 12-AUG-2002; 2002US-00217774.
XX
XX PR 16-AUG-2000; 2000US-0225852P.
XX
XX PR 15-AUG-2001; 2001US-00930872.

PA	(FRID/) FRIDDLE C J.	181	ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAaLeuGlySerSerPro	200
XX	(HILB/) HILBUN E.	541	AGGCCACTTCTTCACACTCTCATGGAACTCGGCAGAGCTGCCCAAGGAGCTCGCCA	600
PI	Friddle CJ, Hilbun E;	201	SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyValAspGluValLeu	220
XX	WPI; 2003-644621/61.	601	TCCCAAGTACTGTACAGAGATCCACAGAGCCCATGCTCTCTGGGGCCAGTGAGTCTG	660
DR	P-PSDB; ADA50760.	221	ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu	240
XX	New polynucleotides encoding human proteins preferably protease sharing	661	GTGACCTCAGAGCATGGAGCTGGCACAATCAACCCCTGCACAGCAGCAGCTTCGCCTG	720
PT	sequence similarity with mammalian proteases, useful for generating	241	GlyLeuProGlnLysHisPheCysGlyArgArgLysLysTyrMetProGlnProPro	260
PT	primers, and probes used to identify drug targets.	721	GGACTGCCACAAAAGCAGCAATTTCTGTGGAAGACGCAAGAAATACATGCCCCAGCCTCCC	780
XX	Disclosure; Page 7-8; 17pp; English.	261	LysGluAspLeuPheLeuLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer	280
CC	The present invention relates to the isolation of novel human protein	781	AAGGAAGACCTCTTTCATCTTCCAGATGAGTATTAAGTCTTGTCTACGCATAAGCGCTCT	840
CC	(NHPs), and the polynucleotide sequences encoding them. The NHPs of the	281	LeuLeuArgSerHisArgAsnGluLeuAsnValGluThrLeuValValValAspLys	300
CC	invention are proteases that share particular homology with zinc	841	CTTCTGAGGTCCCATAGAAATGAAGAACTGAACGTGGAGACCTTGGTGGTGGTGCACAAA	900
CC	metalloproteases. The NHP polynucleotide sequences are useful in gene	301	LysMetMetGlnAsnHisGlyHisGluAsnLeuThrThrTyrValLeuThrLeuLeuAsn	320
CC	therapy techniques for the modulation of NHP expression, for detecting	901	AAGATGATGCAAAACCACTGGCCATGAAATATCACCACCTACGTGCTCCAGTACTCAAC	960
CC	mutant NHPs or inappropriately expressed NHPs for the diagnosis of	321	MetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnLeuAlaLeuVal	340
CC	disease, for screening drugs effective in the treatment of symptomatic or	961	ATGGTATCTCTTATTCAAAGATGGAACAATAGGAGGAAACATCAACATTCGAATTGTA	1020
CC	phenotypic manifestations of perturbing the normal function of NHP in the	341	GlyLeuLeuLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis	360
CC	body. The sequences are also useful for identifying mutations associated	1021	GGTCTGATTCTCTAGAAAGATGAACGCCAGGACTGGTATAGTATCCACCGCAGACCAC	1080
CC	with a particular disease and are useful in prognostic or diagnostic	361	ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis	380
CC	assays. The present sequence encodes a NHP of the invention.	1081	ACCTTAAGTAGCTTCTGCGAGTGGCAGTCTGGATTGATGGGAAAGATGGAGCTCGTCAT	1140
XX	SQ Sequence 1476 BP; 372 A; 393 C; 394 G; 317 T; 0 U; 0 Other;	381	AspHisAlaLeuLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp	400
Alignment Scores:		1141	GACCACGCCATCTTACTGACTGGTCTGGATATATGTTCTTGGAAAGATGAGCCCTGTGAC	1200
Pred. No.:	6.58e-177	401	ThrLeuGlyPheAlaProLysSerGlyMetCysSerLysTyrArgSerCysThrIleAsn	420
Score:	2628.00	1201	ACTTTGGGATTTGCACCCATAAGTGGAAATGTGTAGTAAATATCGCAGCTGCAGATTAA	1260
Percent Similarity:	100.00%	421	GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerClyHisAsnPheGly	440
Best Local Similarity:	100.00%	1261	GAAGATACAGGTCTTGGACTGGCTTCCATATGGCCCTGAGTCTGGACACAACTTTGSC	1320
Query Match:	100.00%	441	MetIleHisAspGlyGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro	460
DB:	9	1321	ATGATTCATGATGGAGAGGGAACATGTGTAAGTCCGAGGGCAACATCATGTGCCCC	1380
US-10-804-457-2 (1-491) x ADA50759 (1-1476)		461	ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis	480
QY	1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaAlaLeuTrpMetLeuLeuAlaGln	1381	ACATTTGGCAGGACGCAATGGAGTCTTCTCTGCTCAGCTCCGCGCCAGTATCTACAC	1440
DB	1 ATGAAGCCCGCGCGCGGATGGCGGCTTGGCGGCTGGATGCTGTTCGGCGAG	481	LysPheLeuArgSerValLysMetProAlaLeu	491
QY	21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaAlaProGlySerPro	1441	AAATTTCTAAGATCAGTGAATAATGCCACTCTC	1473
DB	61 GTGGCCGAGCAGGACCTTGGTGGCCATGGAGCCCGCAGCGGCGGCTGGATG	RESULT 2		
QY	41 SerValProArgProProProAlaGluArgProGlyTrpMetGluLysGlyGluTyr	ABSS7766		
DB	121 AGCGTCCCGCTCTCTCCACCCCGCGAGCGCGCGCTGGATGGAAAGGGCGCAATAT	ID	ABSS7766 standard; cDNA; 1476 BP.	
QY	61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet	XX		
DB	181 GACCTGGTCTCTGCTACGAGTGTGACCACAGGGGCGATACGTGTCCCATGAAATCATG	AC	ABSS7766;	
QY	81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg	XX		
DB	241 CACCATCAGCGCGGAGAGAGCAGTGGCCGTGCCAGGTTGAGTCTCTTCACCTTCGG	DT	05-FEB-2003 (first entry)	
QY	101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerLeuValAla	XX		
DB	301 CTGAAGAGGCTCCAGGCACGACTTCCACGTGGATCTGAGGACTTCCAGCAGCTTAGTGG			
QY	121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro			
DB	361 CTGGCTTTATTTGGCAGAGCTTGGGAAAGACAGGCACTAAGTCTGTGCAGACTTTACCG			
QY	141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla			
DB	421 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGGATCCACAGAACTCTCTCAGTGGCC			
QY	161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluAlaAspTyrPheLeu			
DB	481 CTTTCAACCTGCGCAAGGCTTGTGAGGATGATCGAACAAGAGGAGGAGGAGTACTTCTCTA			

DE cDNA encoding novel human protease #1.
 XX Human; protease; gene therapy; obesity; gene; ss.
 XX OS Homo sapiens.
 XX Key Location/Qualifiers
 FT 1..1476
 FT /*tag= a
 FT /product= "Novel human protease"
 XX
 PN US6448388-B1.
 XX
 PD 10-SEP-2002.
 XX
 PF 15-AUG-2001; 2001US-00930872.
 XX
 PR 16-AUG-2000; 2000US-0225852P.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Friddle CJ, Hilbun E;
 XX
 DR WPI: 2003-074103/07.
 DR P-PSDB; ABG72430.
 XX
 PT New nucleic acid encoding novel human protein (NHP), useful for the
 PT preparation of a medicament for diagnosing, preventing or treating
 PT disorders of diseases associated with the NHP, e.g., obesity.
 XX
 PS Disclosure; Col 13-14; 17pp; English.
 XX
 CC The invention describes a new isolated nucleic acid comprising a sequence
 CC that encodes a fully defined protein sequence comprising 491 or 1224
 CC amino acids or that hybridises under stringent conditions with the 3675-
 CC bp sequence or its complement. The nucleic acid is useful for the
 CC preparation of a medicament for diagnosing, preventing or treating
 CC disorders (e.g. using gene therapy) or diseases associated with the novel
 CC human proteins, e.g., obesity. This sequence encodes a novel human
 CC protease
 XX
 SQ Sequence 1476 BP; 372 A; 393 C; 394 G; 317 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 6.58e-177 Length: 1476
 Score: 2628.00 Matches: 491
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-10-804-457-2 (1-491) x ABS57766 (1-1476)

QY 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaLeuTrpMetLeuLeuAlaGln 20
 DB 1 ATGAAGCCCGCGCGGATGGGGGCTTTGGCGGCGCTGTGGATGCTGTGGCGGAG 60
 QY 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
 DB 61 GTGGCGGAGCAGCACCCTGCTGGCCATGGGACCCGCGGCGGCGCTGGGAGCCGG 120
 QY 41 SerValProArgProProProProAlaGluArgProGlyTrpMetGluLysGlyGlyTr 60
 DB 121 AGCGTCCCGCGTCTCTCCACCGCGGAGCGCGCGGCTGGATGGAAAGGCGCAATAT 180
 QY 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluLeuMet 80
 DB 181 GACCTTGGTCTCTGCTACGAGTTGACCAACGAGGCGGCGATTACGTGTCCCATGAATCATG 240
 QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
 DB 241 CACCATCAGCGCGGAGAGAGAGAGTGGCCGCTGTCCGAGGTTGAGTCTCTTCCACCTTGG 300

QY 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAla 120
 DB 301 CTGAAGGCTCCAGCAGACTTCCACGTGGATCTGAGGACTTCCAGCAGCCCTAGTGCT 360
 QY 121 ProGlyPheLeuValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
 DB 361 CTGGCTTTATTGTGCAGACGTTGGGAAAGACAGGCACTAGTCTGTGAGACTTTACCG 420
 QY 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
 DB 421 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTGGCATCACACAGAAATCTCTCTAGTGGCC 480
 QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu 180
 DB 481 CTTTCAACCTGCCAAGGCTTGTCAAGCATGTATACAAACAGAGAGGAGATTTCTCTTA 540
 QY 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
 DB 541 AGGCCACTTCTTCCACACTCTCATGGAATCTGGCAGAGCTGCCCAAGGCAAGCTTCGCCA 600
 QY 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
 DB 601 TCCACAGTACTGTACAGAGATCCACAGAGCCCATGCTCTCTGGGGCCAGTCAGGTCCTG 660
 QY 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
 DB 661 GTGACCTCAGGACATGGGAGCTGGCACATCAACCCCTGCACAGCAGGACCTTCGCCCTG 720
 QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
 DB 721 GGACTGCCCAAAAGCAGCACTTCTGTGAGAGCAAGAAATACATGCCCCAGGCTCCC 780
 QY 261 LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
 DB 781 AAGGAAGACCTCTTCCATCTTCCAGATGAGTATAAGTCTTGCTTACGGCATAAAGCGCTCT 840
 QY 281 LeuLeuArgSerHisArgAsnGluLeuAsnValGluThrLeuValValAspLys 300
 DB 841 CTTCTAGGTTCCCATAGAAATGAAAGAACTGAACGTGGAGACCTTGTGTGTGTGCACAAA 900
 QY 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrThrTyrValLeuThrIleLeuAsn 320
 DB 901 AAGATGATGCAAAACCATGGCCCATGAAATATCACCACTACGTGCTCAGCATATCTCAAC 960
 QY 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAsnIleVal 340
 DB 961 ATGGTATCTGCTTTATTCAAGATGGAACAATAGGAGGAACAATCAACATTGCAATTGTA 1020
 QY 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisHisAlaSerHis 360
 DB 1021 GGTCTGATTTCTTAGAAGATGAACAGCAGGAGCTGGTGTAAAGTCACCACGACAGCCAC 1080
 QY 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
 DB 1081 ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGATTTGATGGGAAAGATGGGACTGCTCAT 1140
 QY 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
 DB 1141 GACCACGCCATCTTACTGACTGGTCTGGATATATGTTCTTCTGGAAGATGAGCCCTGTGAC 1200
 QY 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
 DB 1201 ACTTTGGGATTTGGCACCATTAAGTGAATGTGTAGTAAATATCGCAGCTGCACGATTAAT 1260
 QY 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
 DB 1261 GAAGATACAGGTCTTTGGAGCTGCCCTTCCACATGCCCCATGAGTCTGGACACAACTTTGGC 1320
 QY 441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
 DB 1321 ATGATTTCATGTGGAGAGGGAACATGTGTAAAGTCCGAGGGAACATCATGTGCTCCCT 1380
 QY 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480

Db 1381 ACATGGCAGGCGCAATGAGTCTTCTCTGTCACCTGCGCGCCAGTATCTACAC 1440
QY 481 LysPheLeuArgSerVallyMetProAlaLeu 491
Db 1441 AATTTCTAAGATCAGTGAATAATGCCAGTCTC 1473

RESULT 3
ID AAL43654
XX AAL43654 standard; DNA; 3675 BP.
AC AAL43654;
XX
DT 05-SEP-2002 (first entry)
DE Human MDT99 protease coding sequence.
XX
KW Human; gene; ds; MDT99 protease; TGF-beta inhibitor;
KW transforming growth factor-beta inhibitor; chronic renal failure.
XX
OS Homo sapiens.

XX Key Location/Qualifiers
FH 1..3675
FT CDS
FT /*tag= a
FT /product= "Human MDT99 protease"

XX WO200251998-A1.
XX
PD 04-JUL-2002.
XX
PF 21-DEC-2001; 2001WO-JP011251.
XX
PR 25-DEC-2000; 2000JP-00393372.
XX
PA (YAMA) YAMANOUCHI PHARM CO LTD.
XX
PI Yamaji N, Nishimura K, Abe K, Ogino M;
XX WPI; 2002-508888/54.
DR P-PSDB; AAO15254.
XX
PT Protease MDT99 and encoded polynucleotide, applicable in diagnosis and
PT screening TGF-beta inhibitors for treatment of chronic renal failure.
XX
PS Example 2; Page 41-48; 60pp; Japanese.
XX
CC The invention comprises the amino acid and coding sequence of the human
CC MDT99 protease. The MDT99 DNA and protein sequences of the invention are
CC useful for screening for transforming growth factor (TGF)-beta inhibitors
CC and for the treatment of chronic renal failure. The present DNA sequence
CC encodes the human MDT99 protease of the invention
XX
SQ Sequence 3675 BP; 873 A; 1007 C; 1020 G; 775 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7.23e-174 Length: 3675
Score: 2592.00 Matches: 486
Percent Similarity: 99.39% Conservative: 3
Best Local Similarity: 98.78% Mismatches: 2
Query Match: 98.63% Indels: 1
DB: 6 Gaps: 0

US-10-804-457-2 (1-491) x AAL43654 (1-3675)
QY 1 MetLysProArgAlaArgGlyTyrArgGlyLeuAlaAlaLeuTrpMetLeuAlaGln 20
Db 1 ATGAAGCCCGCGCGCGATGCGGGCTTGGCGCGCTGTGGATGCTGTGGCGCAG 60
QY 21 ValAlaGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
Db 61 GTGGCCGAGCAGGACCTGTGCGGCATGGGACCGCGCAGCGCGCTGGGAGCCCG 120

QY 41 SerValProArgProProProAlaGluArgProGlyTyrMetGluLysGlyGluTyr 60
Db 121 AGCGTCCGCGTCTCTCCACCGCGAGCGCGCGTGGATGTAAGGCGGCAATAT 180
QY 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluLeuMet 80
Db 181 GACCTGGTCTCTGCGCTACGAGGTGACACAGGGCGGATTAAGTGTCCCATGAATCATG 240
QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
Db 241 CACCATCAGCGCGGAGAGAGAGAGTGGCCGTGTCCGAGGTGGTGTCTCTACCTTCGG 300
QY 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerValAla 120
Db 301 CTGAAGGCTCCAGGCACGACTTCCAGTGTGATCTGAGGACTTCCAGCAGCTAGTGCT 360
QY 121 ProGlyPheLeuValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 361 CTTGGCTTTATTGTGCAGACGTTGGGAAAGACAGGCACCTAAGTCTGTGCAGACTTTACCG 420
QY 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db 421 CCAGAGACTTCTGTTTCTATCAAGGCTCTTTCGATCACACAGAACTCTCTCAGTGGCC 480
QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetLeuArgThrGluGluAlaAspTyrPheLeu 180
Db 481 CTTTCAACCTGCCAAGGCTTGTGAGGCATGATACGAAACAGAAAGGCGAGATTACTTCTTA 540
QY 181 ArgProLeuProSerHisLeuSerTyrLysLeuGlyValArgAlaGlnGlySerSerPro 200
Db 541 AGGCCACTTCTTTCACACCTCTCATGGAACCTGGCAGAGCTGCCAAGGAGGCTGCCA 600
QY 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
Db 601 TCCACGCTACTGTACAAGAGATCCACAGAGCCCATGCTCTCTGGGCGCAGTCTGCTG 660
QY 221 ValThrSerArgThrTyrGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db 661 GTGACCTCAAGGACATGGGAGCTGGCACATCAACCCCTGCACAGCAGCAGCTTCGCGCT 720
QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgGlyLysTyrMetProGlnProPro 260
Db 721 GGAGTCCACAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCGCAGCTCCC 780
QY 261 LysGluAspLeuPheLeuLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db 781 AAGAGAGCTCTTTCATCTTCCAGATGATTAAGTCTTGTCTTACGGCATAAGCGCTCT 840
QY 281 LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValValAspLys 300
Db 841 CTTCTGAGGTCCCATAGAATGAGAACCTGAACTGGAGCTGGAGACCTTGGTGGTTCGACAA 900
QY 301 LysMetMetGlnAsnHisGlyHisGluAsnLeuThrThrTyrValLeuThrLeuAsn 320
Db 901 AAGATGATCAAAACCATGGCCATGAAATAATATCACCCCTAGCTGTCTCACGATCTCAAC 960
QY 321 MetValSerAlaLeuPheLysAspGlyThrLeuGlyGlyAsnLeuAsnLeuAlaLeuVal 340
Db 961 ATGTGATCTGCTTTATTCAAGATGGAAACAATAGGAGGAAACATCAACATGCAATGTGA 1020
QY 341 GlyLeuLeuLeuGluAspGluGlnProGlyLeuValLeuHisHisAlaAspHis 360
Db 1021 GGTCTGATTTCTAGAACATGAACAGCCAGGACTGGTGTATAGTCAACCCAGCAGACCCAC 1080
QY 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 1081 ACCTTAAGTAGTCTTCTGCGAGTGGCAGTCTGGAATGATGGGGAAGATGGGAGCTGCTCAT 1140
QY 381 AspHisAlaLeuLeuLeuThrGlyLeuAspLeuCysSerTyrLysAsnGluProCysAsp 400
Db 1141 GACCAAGCCATCTTACTGACTGCTGGATATATGTTCTCTGGAAAGATGAGGCCCTGTGAC 1200
QY 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420

```
Db 1201 ACTTTGGGATTTGCACCCATAAGTGAATGTAGTAATAATATCGCAGCTGCGACGATTAT 1260
QY 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 1261 GAAGATACAGGCTCTTGGACTGGCCCTTCCACATTGCCCATGAGCTCGGACACAACTTTGGC 1320
QY 441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
Db 1321 ATGATTTCATGATGGAGAGGGAACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCT 1380
QY 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTrpLeuHis 480
Db 1381 ACATTGGCGAGGACGCAATGGAGTCTCTCTGTCACCTGCGAGCGCCGATCTTACAC 1440
QY 481 LysPheLeuArgSer-VallYsMetProAlaLeu 491
Db 1441 AATTTTCTAAGCACCGCTCAAGCTATCTGCCTTG 1474

RESULT 4
ADA50761
ID ADA50761 standard; cdna; 3675 BP.
XX
AC ADA50761;
XX
DT 20-NOV-2003 (first entry)
XX
DE cDNA encoding novel human protein (NHP) #2.
XX
KW Novel human protein; NHP; protease; zinc metalloproteinase; gene therapy;
KW human; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3675
FT /tag= a
FT /product= "NHP #2"
XX
PN US2002193583-A1.
XX
PD 19-DEC-2002.
XX
PF 12-AUG-2002; 2002US-00217774.
XX
PR 16-AUG-2000; 2000US-0225852P.
XX
PR 15-AUG-2001; 2001US-00930872.
XX
XX (FRID/) FRIDDLE C J.
XX (HILB/) HILBUN E.
XX
PI Friddle CJ, Hilbun B;
XX
XX WPI; 2003-644621/61.
XX DR P-PSDB; ADA50762.
XX
XX New polynucleotides encoding human proteins preferably protease sharing
XX sequence similarity with mammalian proteases, useful for generating
XX primers, and probes used to identify drug targets.
XX
XX Claim 2; Page 9-11; 17pp; English.
XX
XX The present invention relates to the isolation of novel human protein
XX (NHPs), and the polynucleotide sequences encoding them. The NHPs of the
XX invention are proteases that share particular homology with zinc
XX metalloproteinases. The NHP polynucleotide sequences are useful in gene
XX therapy techniques for the modulation of NHP expression, for detecting
XX mutant NHPs or inappropriately expressed NHPs for the diagnosis of
XX disease, for screening drugs effective in the treatment of symptomatic or
XX phenotypic manifestations of perturbing the normal function of NHP in the
XX body. The sequences are also useful for identifying mutations associated
XX with a particular disease and are useful in prognostic or diagnostic
XX assays. The present sequence encodes a NHP of the invention.
```

```
XX SQ Sequence 3675 BP; 873 A; 1008 C; 1020 G; 774 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 7,23e-174 Length: 3675
Score: 2592.00 Matches: 486
Percent Similarity: 99.39% Conserv: 3
Best Local Similarity: 98.78% Mismatches: 2
Query Match: 98.63% Indels: 1
DB: 9 Gaps: 0
US-10-804-457-2 (1-491) x ADA50761 (1-3675)
QY 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaLeuTrpMetLeuLeuAlaGln 20
Db 1 ATGAAGCCCGCGCGCGGATGGCGGGCTTGGCGGCTGTGGATGTGTGGCGCAG 60
QY 21 VallAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAalaAalaProGlySerPro 40
Db 61 GTGGCCGAGCAGGACACCTGCGTGCCTATGGACCCCGCAGCGGAGCGCTGGGAGCCG 120
QY 41 SerValProArgProProProAlaGluArgProGlyTrpMetGluLysGlyGlyTyr 60
Db 121 AGGTCCCGCGTCTCTCTCCACCCGCGGAGCGCGCGGCTGGATGGATGGAAGGCGGATAT 180
QY 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
Db 181 GACCTGGTCTCTGCTACGAGTTGACCCACAGGCGGATACGTGTCTCCATGAAATCATG 240
QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
Db 241 CACCATCAGCGCGGAGAGAGCAGTGGCGGTGTCGAGGTGTAGTCTCTTCCACTTCGG 300
QY 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerValAla 120
Db 301 CTGAAGGCTCTCAGGACGACATTCACGCTGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360
QY 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 361 CTGGCTTTATTGTGACAGCTTGGGAAAGACAGGCACTAAGTCTGTGCAGACTTTACCG 420
QY 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db 421 CCAGAGGACTTCTGTCTTATCAAGGCTTTTGGCATCACACAGAACTCTCCAGTGGCC 480
QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu 180
Db 481 CTTTCACTGCGCAAGGCTTGTCCAGCATGATACGACAGAGAGGCGAGATTACTTCTTA 540
QY 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
Db 541 AGGCCACTTCTTTCACACCTCTCATGGAAACTGGCGAGAGCTGCCCAAGGCAAGCTCGCCA 600
QY 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
Db 601 TCCACGCTACTGTACAGAGATCCACAGAGCCCCATGCTCTCTGGGGGCCAGTGGAGTCTG 660
QY 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db 661 GTGACCTCAAGGACATGGAGCTGGCACAATCAACCCCTGCACAGCAGCACTTTCGCTG 720
QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
Db 721 GGACTGCCACAAAGCAGCACTTCTGTGGAAGACCAAGAAATACATGCCCCAGCTCTCC 780
QY 261 LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db 781 AAGGAAGACCTCTTTCATCTTCCAGATGAGTATAGTCTTGTCTACGGCATAGGGCTCT 840
QY 281 LeuLeuArgSerHisArgAsnGluLeuLeuValGluThrLeuValValAspLys 300
Db 841 CTTCTGAGGTCCCATAGAAATGAAGAACTGAACGTGGAGACCTTGGTGTGTGCACAA 900
```


QY 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrThrTyrValLeuThrIleLeuAen 320
 Db |||||
 QY 901 AAGATGATGCAAAACCATGGCCATGAATATATACCACTACCTGTCTCAGGATACCAAC 960
 Db |||||
 QY 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyAsnIleAlaIleVal 340
 Db |||||
 QY 961 ATGGTATCTGCTTATTTCAAAGATGGAACAATAGGAGGAACAATCAATTCGAATTGTA 1020
 Db |||||
 QY 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisIleAlaPheHis 360
 Db |||||
 QY 1021 GGCTCTGATCTCTTAGAAGATGAACAGCCAGGACTGGTATAGTACCAACAGCAGCAC 1080
 Db |||||
 QY 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
 Db |||||
 QY 1081 ACCTTAAGTAGCTCTGCCAGTGGCAGTCTGGATTGATGGGGAAGATGGGACTCGTCAT 1140
 Db |||||
 QY 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
 Db |||||
 QY 1141 GACCAGCCATCTTACTGACTGGTCTGGATATATGTTCTGGAAGATGAGCCCTGTGAC 1200
 Db |||||
 QY 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAen 420
 Db |||||
 QY 1201 ACTTTGGGATTTGCCACCATAGTGAATGTGTAGTAAATATCGCAGCTGCACGATTAA 1260
 Db |||||
 QY 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
 Db |||||
 QY 1261 GAAGATACAGGTCTTTGGACTGGCCTTCACCATGGCCATGAGTCTGGACACAACTTTGCG 1320
 Db |||||
 QY 441 MetIleHisAspGlyGluGlyAsnMetCysValLysSerGluGlyAsnIleMetSerPro 460
 Db |||||
 QY 1321 ATGATTCATGATGGAGAGGGAAACATGTGTAAAGTCCGAGGGGAACATCATGTGCCCT 1380
 Db |||||
 QY 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
 Db |||||
 QY 1381 ACATTGGCAGGACGCAATGGAGTCTTCTCTGTCACCTCGAGCCGCGCAGTATCTACAC 1440
 Db |||||
 QY 481 LysPheLeuArgSerValLysMetProAlaLeu 491
 Db |||||
 QY 1441 AAATTTCTAAGACCGCTCAAGCTATCTGCCTTG 1474
 Db |||||
 RESULT 5
 ID ABS57767
 AC ABS57767 standard; cDNA; 3675 BP.
 AC ABS57767;
 XX 05-FEB-2003 (first entry)
 DT Novel human protease coding sequence.
 DE Human; protease; gene therapy; obesity; gene; ss.
 KW Homo sapiens.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..3675
 FT /*tag= a
 FT /product= "Novel human protease"
 XX US6448388-B1.
 XX 10-SEP-2002.
 XX 15-AUG-2001; 2001US-00930872.
 XX 16-AUG-2000; 2000US-0225852P.
 XX (LEXI-) LEXICON GENETICS INC.
 XX Fziddle CJ, Hilbun E;
 XX WPI; 2003-074103/07.
 DR P-PSDB; ABG72431.

XX
 PT New nucleic acid encoding novel human protein (NHP), useful for the
 PT preparation of a medicament for diagnosing, preventing or treating
 PT disorders or diseases associated with the NHP, e.g., obesity.
 XX
 PS Claim 2; Col 17-20; 17pp; English.
 XX
 CC The invention describes a new isolated nucleic acid comprising a sequence
 CC that encodes a fully defined protein sequence comprising 491 or 1224
 CC amino acids or that hybridises under stringent conditions with the 3675-
 CC bp sequence or its complement. The nucleic acid is useful for the
 CC preparation of a medicament for diagnosing, preventing or treating
 CC disorders (e.g. using gene therapy) or diseases associated with the novel
 CC human proteins, e.g., obesity. This sequence encodes a novel human
 CC protease
 XX
 SQ Sequence 3675 BP; 873 A; 1008 C; 1020 G; 774 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 7,23e-174 Length: 3675
 Score: 2592.00 Matches: 486
 Percent Similarity: 99.39% Conservative: 3
 Best Local Similarity: 98.78% Mismatches: 2
 Query Match: 98.63% Indels: 1
 DB: 10 Gaps: 0
 US-10-804-457-2 (1-491) x ABS57767 (1-3675)
 QY 1 MetLysProArgAlaArgGlyTyrArgGlyLeuAlaAlaLeuTrpMetLeuLeuAlaGln 20
 Db |||||
 QY 1 ATGAAGCCCGCGCGCGCGATGGCGGGCTTGGCGGGCTGTGGATGCTGTGGCCGAG 60
 Db |||||
 QY 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
 Db |||||
 QY 61 GTGGCCGAGCAGACACCTGCGTGGCCATGGGACCGCGCAGCGCGCTGGAGCCCG 120
 Db |||||
 QY 41 SerValProArgProProProAlaGluArgProGlyTyrMetGlyLysGlyGlyTyr 60
 Db |||||
 QY 121 AGGTCCCGGTCTCTCTCCACCGCGAGCGCGGGCTGGATGGAAGGCGGAATAT 180
 Db |||||
 QY 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
 Db |||||
 QY 181 GACCTGGTCTCTGCTTACGAGGTGACACAGGGCGGATTTACGTGCTCCCATGAATCATG 240
 Db |||||
 QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
 Db |||||
 QY 241 CACCATCAGCGCGGAG 300
 Db |||||
 QY 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAla 120
 Db |||||
 QY 301 CTGAAGGGCTCCAGGCACGACTTCCACGTGGATCTCAGGACTTCCAGAGCCTAGTGGCT 360
 Db |||||
 QY 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
 Db |||||
 QY 361 CCTGGCTTTATTGTGCAGACGTTGGGAAACAGACAGGCACCTAAGTCTGTGCAGACTT 420
 Db |||||
 QY 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
 Db |||||
 QY 421 CCAGAGGACTTCTGTCTTCTATCAAGGCTCTTTCGATCACAGAACTCTCAGTGGGC 480
 Db |||||
 QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu 180
 Db |||||
 QY 481 CTTTCAACCTGCAAGGGCTTGTCCAGCATGATACGAACAGAGAGAGAGAGAGAGAGAG 540
 Db |||||
 QY 181 ArgProLeuProSerHisIleLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
 Db |||||
 QY 541 AGGCCACTTCTTTCACACCTCTCATGGAAACTCGGAGAGCTGCCCAAGGAGCTCGCCA 600
 Db |||||
 QY 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyValAsnGluValLeu 220
 Db |||||
 QY 601 TCCACGCTACTGTACAGAGATCCACAGAGCCCATGCTCTCTGGGGCCAGCTGAGTCTCTG 660
 Db |||||
 QY 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240

```
Db 661 GTGACCTCAGAGCATGGAGCTGCACATCAACCCCTGCACGAGGACCTTCGCCTG 720
Qy 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnPro 260
Db 721 GGACTGCCACAAAAGCAGCATTTCTGTGGAGAGCGCAAGAAATACATGCCCGCCTCC 780
Qy 261 LysGluAspLeuPheIleLeuProAspGlnTyrLysSerCysLeuArgHisLysArgSer 280
Db 781 AAGGAAGACCTCTTCATCTTGCAGATGATGATAAGTCTTGCTTACGGCATAAGCGCTC 840
Qy 281 LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValAspLys 300
Db 841 CTTCTGAGGTCCCATAGAAATAGAACTGAACGTGGAGACCTTGCTGGTGGTGCACAAA 900
Qy 301 LysMetGlnAsnHisGlyHisGluAsnIleThrThrValLeuThrIleLeuAsn 320
Db 901 AAGATGATGCAAAACCATGGCCATGAAATATACACCACTAGTGTCTCAGCATCTCAAC 960
Qy 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyLysAsnIleAsnIleAlaVal 340
Db 961 ATGGTATCTGCTTTATTCAAGATGGAACAAATAGGAGGAAACATCAACATTCGCAATTGTA 1020
Qy 341 GlyLeuIleLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis 360
Db 1021 GGTCTGATCTTCTAGAAGATGAACGACGAGCTGGGTGATAAGTCAACGACGACCCAC 1080
Qy 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 1081 ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGGAAGATGGGACTGTCTAT 1140
Qy 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db 1141 GACCAAGCCATCTTACTGACTGCTTGGATATATCTCTCGAAGAAATGAGCCCTGTGAC 1200
Qy 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
Db 1201 ACTTTGGGATTTGACACCATAGTGGAAATGTGTAGTAAATATCGCAGCTGCACGATTAA 1260
Qy 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 1261 GAAGATACAGGCTCTGGACTGGCTTACCATTTGCCATGAGTCTGGACACACTTTGGC 1320
Qy 441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
Db 1321 ATGATTTCATGATGGAGAGGGAACATGTGTAAAGTCCGAGGCAACATCATGTGCCCT 1380
Qy 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
Db 1381 ACATTGGCAGGACGCAATGGAGTCTTCTCCTGGTCACCCCTGCAGCCGCGATATCTACAC 1440
Qy 481 LysPheLeuArgSer-ValLysMetProAlaLeu 491
Db 1441 AATTTCTAAGCACCGCTCAGATATCTGCCTTG 1474
```

RESULT 6

ACC85689

ID ACC85689 standard; DNA; 3675 BP.

AC ACC85689;

XX ACC85689;

XX 22-APR-2004 (first entry)

XX Human protease MDTs9 related coding sequence.

DE DE

XX Promoter; human; protease; MDTs9; chronic kidney failure; nephrotropic;

KW gene; ds.

XX OS

XX Homo sapiens.

XX FN

XX WO2004001047-A1.

XX PD

XX 31-DEC-2003.

```
XX 19-JUN-2003; 2003WO-JP007807.
XX 20-JUN-2002; 2002JP-00180543.
XX (YAMA ) YAMANOUCHI PHARM CO LTD.
XX Takeda M, Abe K, Yamaji N;
XX WPI; 2004-071739/07.
XX P-PSDB; ABM79745.
XX New transforming growth factor-beta derived protease MDTs 9 promoter, for
XX screening remedies and/or preventives for chronic kidney failure.
XX Disclosure; Page 30-39; Opp; Japanese.
XX The present invention relates to a polynucleotide, which has a promoter
XX activity. The sequence is derived from the protease MDTs9 gene. The
XX promoter is useful in screening remedies and/or preventives for chronic
XX kidney failure. The present sequence is a polynucleotide shown in the
XX exemplification of the invention
XX SQ Sequence 3675 BP; 873 A; 1007 C; 1020 G; 775 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 7.23e-174 Length: 3675
Score: 2592.00 Matches: 486
Percent Similarity: 99.39% Conservative: 3
Best Local Similarity: 98.78% Mismatches: 2
Query Match: 98.63% Indels: 1
DB: 12 Gaps: 0
US-10-804-457-2 (1-491) x ACC85689 (1-3675)
```

```
Qy 1 MetLysProArgAlaAArgGlyTrpArgGlyLeuAlaAAlaLeuTrpMetLeuAlaGln 20
Db 1 ATGAAGCCCGCGCGCGGATGGCGGCTTGGCGGCTGTGGATGCTGTGTGGCGGAG 60
Qy 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaAAlaProGlySerPro 40
Db 61 GTGGCCGAGCAGGACCTCGTGGCCATGGGACCCGCGGAGCGGCGCTGGAGCCCG 120
Qy 41 SerValProArgProProProAlaGluArgProGlyTrpMetGluLysGlyGluTyr 60
Db 121 AGCGTCCCGCTCTCTCTCCACCCGCGGAGCGGCGGCTGGATGCGAAAGCGCAATAT 180
Qy 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
Db 181 GACCTGGTCTCTGCTTACGAGGTGACCAAGGGGCGATTACGTGCTCCATGAAATCATG 240
Qy 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
Db 241 CACCATCAGCGCGGAGAGAGCAGTGGCCGTGTCCGAGGTGAGTCTCTTCCACTTCG 300
Qy 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAla 120
Db 301 CTGAAAGGCTCCAGCACGACTTCCACGTGGATCTGAGGACTTCCAGCAGCGCTAGTGGCT 360
Qy 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 361 CCTGGCTTTATTGTGCAGACGTGGGAAAGACAGCAGCTAAGTCTGTGCAGACTTACCG 420
Qy 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db 421 CCAGAGGACTTCTGTCTTATCAGGCTCTTTGGATCAGACAGAAATCTCTCAGTGGCC 480
Qy 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluAlaAspTyrPheLeu 180
Db 481 CTTTCACCTGCCAAGGCTTGTCTAGGCATGATACGACAGAGGAGGAGGAGGAGTACTTCTA 540
Qy 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
```

```
Db 541 AGCCACTTCCTTCAACCTCTCATGGAACTCGCAGAGCTGCCAAGGAGCTCGCCA 600
Qy SerHisValLeuTyrIleArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
Db 601 TCCACGTAAGTGTACAGAGATCCACAGAGCCCATGCTCTCTGGGGCCAGTGAGGTCTG 660
Qy 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db 661 GTGACCTCAAGGACATGGAGCTGGCAGCATCAACCCCTGCGACAGCGACCTTCGGCTG 720
Qy 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
Db 721 GACCTGCCACAAAGACAGATTCGTGGAAAGAGCGCAAGAAATACATGCCCGACCTCC 780
Qy 261 LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db 781 AAGGAAGACCTCTTCATCTTCGAGATGAGTATAGTCTTGCTTACGGCATAGCGCTCT 840
Qy 281 LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValAspLys 300
Db 841 CTTCTGAGGTCCCATAGAAATGAAGAACTGAAGCTGAGACCTTGGTGGTTCGACAA 900
Qy 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrTyrValLeuThrIleLeuAsn 320
Db 901 AAGATGATGCAAAACCATGCCATGAANAATATCACCACCTACGTGCTCAGTACTCAAC 960
Qy 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyLysAsnIleAlaIleVal 340
Db 961 ATGGTATCTCTCTTATTCAAAGATGGAACAATAGGAGGAACAATCAACATTCGATTGTA 1020
Qy 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis 360
Db 1021 GGTCTGATCTCTAGAGATGAGACAGCCAGGACTGGTGTATAGTCAAGTCAACAGCAC 1080
Qy 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 1081 ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCT 1140
Qy 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db 1141 GACCACGCCATCTTACTGACTGTCTGGATATATGTTCTCGAAGATGAGCCCTGTGAC 1200
Qy 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
Db 1201 ACTTTGGGATTTGGCACCACCAATAGTGAATGTGTAGTAATATCCAGCTGCACCAATAT 1260
Qy 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 1261 GAAGATACAGTCTTGGAGCTGGCCCTTACCATTGCCATTGAGTCTGGACACAACTTTGGC 1320
Qy 441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
Db 1321 ATGATTTCATGTGGAGAGGGAACATGTGTAAAGTCCGAGGGCAACATCATGTCCCT 1380
Qy 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
Db 1381 ACATGGGAGGAGCGCAATGGAGTCTTCTCTGTGTACCTGTGACCGCCAGTATCTACAC 1440
Qy 481 LysPheLeuArgSerValLysMetProAlaLeu 491
Db 1441 AAATTTCTAAGCACCGCTCAAGCTATCTGCTCTG 1474
RESULT 7
ADA50763
ID ADA50763 standard; DNA, 4042 BP.
XX
AC ADA50763;
XX
20-NOV-2003 (first entry)
XX
DT Novel human protein (NHP) ORF and flanking sequence.
XX
DE Novel human protein; NHP; protease; zinc metalloproteinase; gene therapy;
XX
KW
```

```
human; gene; ds.
XX
OS Homo sapiens.
XX
PN US2002193583-A1.
XX
19-DEC-2002.
XX
12-AUG-2002; 2002US-00217774.
XX
16-AUG-2000; 2000US-0225852P.
XX
15-AUG-2001; 2001US-00930872.
XX
(FRID/) FRIDDLE C J.
PA (HILB/) HILBUN E.
XX
Fridde CJ, Hilbun E;
XX
WPI; 2003-644621/61.
XX
New polynucleotides encoding human proteins preferably protease sharing
sequence similarity with mammalian proteases, useful for generating
primers, and probes used to identify drug targets.
XX
Disclosure; Page 14-16; 17pp; English.
XX
The present invention relates to the isolation of novel human protein
(NHPs), and the polynucleotide sequences encoding them. The NHPs of the
invention are proteases that share particular homology with zinc
metalloproteinases. The NHP polynucleotide sequences are useful in gene
therapy techniques for the modulation of NHP expression, for detecting
mutant NHPs or inappropriately expressed NHPs for the diagnosis of
disease, for screening drugs effective in the treatment of symptomatic or
phenotypic manifestations of perturbing the normal function of NHP in the
body. The sequences are also useful for identifying mutations associated
with a particular disease and are useful in prognostic or diagnostic
assays. The present sequence represents a NHP ORF and flanking sequences.
XX
SQ Sequence 4042 BP; 947 A; 1117 C; 1130 G; 848 T; 0 U; 0 Other;
```

Alignment Scores:

```
Pred. No.: 8.13e-174 Length: 4042
Score: 2592.00 Matches: 486
Percent Similarity: 99.39% Conservative: 3
Best Local Similarity: 98.78% Mismatches: 2
Query Match: 98.63% Indels: 1
DB: Gaps: 0
```

US-10-804-457-2 (1-491) x ADA50763 (1-4042)

```
Qy 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaAlaLeuTrpMetLeuLeuAlaGln 20
Db 99 ATGAAGCCCCCGCGCGGATGGCGGCTTGGCGGCTGTGGATGCTGCTGGCGAG 158
Qy 21 VallalaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAaProGlySerPro 40
Db 159 GTGGCCGAGCAGGACCTCGTGGCCATGGGACCCGCGGCGAGCGCTTGGAGCCCC 218
Qy 41 SerValProArgProProProAlaGluArgProGlyTrpMetGluLysGlyTyr 60
Db 219 AGCTTCGGGTCTCTCTCCACCGGAGCGGCGGCTGGATGGAAGGCGGAATAT 278
Qy 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIleMet 80
Db 279 GACCTGGTCTCTGCTACGAGGTGGACCAAGGCGGATGCTGCTCCATGAATCATG 338
Qy 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
Db 339 CACCATCAGCGCGGAGAGAGAGAGTGGCGGTGTCCGAGGTTGAGTCTCTTACCTCGG 398
Qy 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerLeuValAla 120
Db 399 CTGAAAGGCTCCAGGACGACTTCCACGCTGGATCTGAGGACTTCCAGGCGCTAGTGGCT 458
```

Qy 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 459 CTTGGCTTTATTGTGCAGACGTTGGGAAGAAGCAGGCACTAAGTCTGTGCAGACTTTACCG 518
Qy 141 ProGluAspPheCysPheTyTGlnGlySerLeuArgSerHisAsnSerSerValAla 160
Db 519 CAGAGGAGCTTCTGTTTCTATCAGGCTCTTTGGCATCACAGAAACTCTCTCAGTGGCC 578
Qy 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu 180
Db 579 CTTTCAACCTGCCAAGGCTTGTTCAGGCATGATACAGACAGAGAGCGAGATTCTTCTTA 638
Qy 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
Db 639 AGGCCACTTCTCTCACCTCTCATGGAACTCGGCAGAGCTGCCCAAGCGCAGCTCGCCA 698
Qy 201 SerHisValLeuTyLysArgSerThrGluProHisAlaProGluValSerGluValLeu 220
Db 699 TCCCAACGTACTGTACAAGAGATCCACAGAGCCCCATGCTCTCTGGGGCCAGTGAGTCCCTG 758
Qy 221 ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db 759 GTGACTCTCAGACATGGAGCTGGCAATCAATCAACCTTCACAGCAGCAGCTTCGCTG 818
Qy 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
Db 819 GGACTGCCCAAAAGCAGCACTTCTGTGGAAGCAGCAAGAAATACATGCCCGAGCTCC 878
Qy 261 LysGluAspLeuPheIleLeuProAspGlnTyrLysSerCysLeuArgHisLysArgSer 280
Db 879 AAGGAAGACCTCTTCATCTTGCACATGAGTAAAGTCTTGTCTACGGCATAAGCGCTCT 938
Qy 281 LeuLeuArgSerHisArgAsnGluGluLeuValGluThrLeuValValAspLys 300
Db 939 CTTCTGAGTCCCATAGAAATGAGAACTGAACGTGGAGACCTTGGTGGTGCACAAA 998
Qy 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrTyrValLeuThrIleLeuAsn 320
Db 999 AAGATGATGCAAAACCATGGCCATGAAATATACCACTACCTAGTCTCAGCATACTCAAC 1058
Qy 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyLysAsnIleAlaIleVal 340
Db 1059 ATGGTATCTGCTTTATTCAAGATGGAACAATAGAGGAAACATCAATGCAATTGTA 1118
Qy 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis 360
Db 1119 GGTCTGATTTCTTAGAAGATGAACAGCAGGACTGGTGAAGTCAACACGACGACAC 1178
Qy 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 1179 ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCA 1238
Qy 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db 1239 GACCACGCCATCTTACTGACTGCTGTGGATATATGTTCTGTGAAGAAATGAGCCCTGTGAC 1298
Qy 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
Db 1299 ACTTTGGGATTTGCACCCATAGTGAATGTGTAGTAATATCGCAGCTGCACGATTAT 1358
Qy 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 1359 GAAGATACAGGCTTGGAGCTGGCCTTCACCATTTGCCATGCTGGACACAACTTTGGC 1418
Qy 441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
Db 1419 ATGATTCTATGTGGAAGGAACTGTGTAAAGTCCGAGGGCAACATCATGTCTCCCT 1478
Qy 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
Db 1479 ACATTTGGCAGGACGAATGAGTCTTCTCTCTGTCACCTGACCGCCGAGTATCTACAC 1538

Qy 481 LysPheLeuArgSer-VallysMetProAlaIleu 491
Db 1539 AAATTTCTAAGCACCCTCAAGCTATCTGCCTTG 1572
RESULT 8
ABS57768
ID ABS57768 standard; cDNA; 4042 BP.
XX
AC ABS57768;
XX
DT 05-FEB-2003 (first entry)
XX
DE cDNA encoding novel human protease #2.
XX
KW Human; protease; gene therapy; obesity; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 99..3773
FT /*tag= a
FT /product= "Novel human protease"
XX
PN US6448388-B1.
XX
PD 10-SEP-2002.
XX
PF 15-AUG-2001; 2001US-00930872.
XX
PR 16-AUG-2000; 2000US-0225852P.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Friddle CJ, Hilbun E;
XX
DR WPI; 2003-074103/07.
DR P-FSDB; ABG72431.
XX
PT New nucleic acid encoding novel human protein (NHP), useful for the
PT preparation of a medicament for diagnosing, preventing or treating
PT disorders or diseases associated with the NHP, e.g., obesity.
XX
PS Disclosure; Col 25-30; 17pp; English.
XX
CC The invention describes a new isolated nucleic acid comprising a sequence
CC that encodes a fully defined protein sequence comprising 491 or 1224
CC amino acids or that hybridises under stringent conditions with the 3675-
CC bp sequence or its complement. The nucleic acid is useful for the
CC preparation of a medicament for diagnosing, preventing or treating
CC disorders (e.g. using gene therapy) or diseases associated with the novel
CC human proteins, e.g., obesity. This sequence encodes a novel human
CC protease
XX
SQ Sequence 4042 BP; 947 A; 1117 C; 1130 G; 848 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8,13e-174 Length: 4042
Score: 2592.00 Matches: 486
Percent Similarity: 99.39% Conservative: 3
Best Local Similarity: 98.78% Mismatches: 2
Query Match: 98.63% Indels: 1
DB: 10 Gaps: 0
US-10-804-457-2 (1-491) x ABS57768 (1-4042)

Qy 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaAlaLeuTrpMetLeuLeuAlaGln 20
Db 99 ATGAAGCCCGCGCGCGGATGGCGGCTTGGCGGCTGTGGATGCTGTGCGCGAG 158
Qy 21 ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
Db 159 GTGGCCGAGCAGGACCTTGGTGGCCATGGACCCGAGCGGCGCTGGAGCCCG 218

QY 41 SerValProArgProProProProAlaGluArgProGlyTyrMetGluLysGlyGluTyr 60
Db 219 AGCGTCCCGGCTCTCTCCACCCGCGAGCGCGGCTGGATGGAAAGGGCGAATAT 278
QY 61 AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluLeuMet 80
Db 279 GACCTGGTCTCTGCTACGAGTTGACACAGGGCGGATACGTGTCCATGAATCATG 338
QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
Db 339 CACCATCAGCGCGGAGAGAGAGTGGCGCTGTCGAGGTTGAGTCTCTTACCTTCGG 398
QY 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerLeuValAla 120
Db 399 CTGAAGGCTCCAGCAGCACTTCCAGCTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 458
QY 121 ProGlyPheLeuValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
Db 459 CTTGGCTTATGTGCGAGCGTTGGGAAGACAGGCACTAAGTCTGTGCGAGACTTTACCG 518
QY 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
Db 519 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTCGGATCAGACAGAACTCTTCAGTGGCC 578
QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluAlaAlaAspTyrPheLeu 180
Db 579 CTTTCAACCTGCCAGGCTTGTGAGCATGATACGAACAGAGGAGGACGATTTACTTCTTA 638
QY 181 ArgProLeuProSerHisLeuSerTyrLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
Db 639 AGGCCACTTCTTACACCTCTCATCGAAGCTCGGCGAGGCTGCCCAAGGCGAGCTCGCCA 698
QY 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
Db 699 TCCAGCTACTGTACAAAGAGATCCAGAGCCCATGCTCTCTGGGCGCCAGTGAGTCTGT 758
QY 221 ValThrSerArgThrTyrGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
Db 759 GTGACCTCAAGGACATGGAGCTGGCATCAACCTTGCACAGCAGGACCTTCGCGCTG 818
QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
Db 819 GGACTGCCCAAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCCAGCTTCCC 878
QY 261 LysGluAspLeuPheLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db 879 AAGGAAGACCTTCTTATCTGCCAGATGAGTATAGTCTTGTCTACGGCATAAGCGCTCT 938
QY 281 LeuLeuArgSerHisArgAsnGluLeuAsnValGluThrLeuValValValAspLys 300
Db 939 CTTCTGAGGTCCCATAGAAATGAAGAACTGAACGTGGAGACCTTGTGTGGTTCGACAAA 998
QY 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrTyrValLeuThrIleLeuAsn 320
Db 999 AAGATGATGCAAAACCATGGCCATGGAATAATATACACCTACGTGTCTACGATCTCAAC 1058
QY 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAsnIleAlaIleVal 340
Db 1059 ATGGTATCTGCTTTATTCAAAGATGGAAATAATAGGAGGAACATCAACATTCGAATGTA 1118
QY 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis 360
Db 1119 GGTCTGATTTCTTAGAAGATGAACAGCCAGGACTGGTATAGTCAACACGACAGACCAC 1178
QY 361 ThrLeuSerSerPheCysGlnTyrGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 1179 ACCTTAGTAGTCTTCCAGTGGCAGTCTGGATTGATGGGAAGATGGGACTCGTCAT 1238
QY 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTyrLysAsnGluProCysAsp 400
Db 1239 GACCAGCCCATCTTACTGCTGTGGATATATGTTCTGGAAGAAATGAGCCCTGTGAC 1298
QY 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420

Db 1299 ACTTTGGGATTTGCCACCCCATAGTGAATGTGTAGTAATAATATCGCAGCTGCACGATTAT 1358
QY 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 1359 GAAGATACAGGTCTTTGGACTGGCTTCCACATGCCCATGAGTCTGGACACAACATTTGGC 1418
QY 441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
Db 1419 ATGATTCATGATGGGAAGGGAACATGTGTAAAAAGTCCGAGGCGCAACATCATGTCCCT 1478
QY 461 ThrLeuAlaGlyArgAsnGlyValPheSerTyrSerProCysSerArgGlnTyrIleHis 480
Db 1479 ACATTTGGCAGCAGCATGAGTCTTCTCTGGTCACTTCGCGCCGCGCAGTATCTACAC 1538
QY 481 LysPheLeuArgSerValLysMetProAlaLeu 491
Db 1539 AAATTTCTAAGCACCGCTCAAGCTATCTGCGCTTG 1572

RESULT 9
ADRO7547
ID ADRO7547 standard; cDNA; 3287 BP.
XX
AC ADRO7547;
XX
DT 04-NOV-2004 (first entry)
XX
DE Full length human cDNA useful for treating neurological disease Seq 1053.
XX
KW gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;
KW osteoporosis; neurological disease; Alzheimer's disease;
KW Parkinson's disease; dementia; short memory; cancer;
KW sense or motor function; emotional reaction; fear response; panic;
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cyostatic;
KW tranquiliser.
XX
OS Homo sapiens.
XX
XX
XX EP1447413-A2.
XX
PD 18-AUG-2004.
XX
PF 12-FEB-2004; 2004BP-00003145.
XX
PR 14-FEB-2003; 2003JP-00102207.
PR 09-MAY-2003; 2003JP-00131452.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T,
PI Wakamatsu A, Ishii S, Nagai K, Irie R;
XX
DR WPI: 2004-583265/57.
DR P-PSDB; ADR09503.
XX
PT New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS Claim 1; SEQ ID NO 1053; 2686pp; English.
XX
CC This invention relates to novel, isolated full length human cDNA
CC molecules and the encoded proteins thereof. Specifically, it refers to
CC cDNA clones obtained by an oligo-capping method, where none of these
CC clones are identical to any known human mRNAs. The present invention
CC describes an immunoassay to identify agonists and antagonists, as well as
CC antibodies, antisense molecules and siRNAs that can all be used to bind
CC to and modulate expression of the cDNA molecules. As such, these
CC molecules are useful for diagnostic markers or therapeutic targets for
CC the various diseases or morbid states. In particular, they are useful in
CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
CC disease, Parkinson's disease, dementia, short memory and various cancers,
CC as well as for maintaining equilibrium of sense or motor function, and
CC for treating emotional reaction, fear response and panic. Accordingly,

CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian, cytoskeletal and translational activities. This polynucleotide is a full length human cDNA sequence of the invention. NOTE: This sequence is not given in the sequence listing of the specification but can be obtained on CC CD-ROM from the European Patent Office, Vienna Sub-office.

XX
SQ Sequence 3287 BP; 789 A; 844 C; 877 G; 777 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.43e-173 Length: 3287
Score: 2587.00 Matches: 485
Percent Similarity: 99.10% Conservative: 3
Best Local Similarity: 98.58% Mismatches: 3
Query Match: 13 Indels: 1
DB: 13 Gaps: 0

US-10-804-457-2 (1-491) x ADR07547 (1-3287)

QY 1 MetLysProArgAlaArgGlyTrpArgGlyLeuAlaLeuTrpMetLeuLeuAlaGln 20
DB 139 ATGAAGCCCGCGCGGATGGCGGCTTGGCGGCGCTGTGGATGCTGTGGCGCAG 198
QY 21 ValAlaGluAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro 40
DB 139 GTGGCGGAGCAGCAGCTGCTGGCGCATGGACCGCGCGCGCTGGAGCCCG 258
QY 41 SerValProArgProProProAlaGluArgProGlyTrpMetGlyLysGlyGluTrp 60
DB 259 AGCGTCCCGGCTCTCCACCGCGGAGCGCGGCTGGATGGAAAGCGCGAATAT 318
QY 61 AspLeuValSerAlaTrpGluValAlaPheHisArgGlyAspTrpValSerHisGluLeuMet 80
DB 319 GACCTGGTCTCTGCTTACGAGGTGACACACAGGGCGATTACGTGTCCCATGAATCATG 378
QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
DB 379 CACCATCAGCGCGGAGAGAGCAGTGGCCGTCCGAGGTGAGTCTCTTCCACCTTGG 438
QY 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgTrpSerSerLeuValAla 120
DB 439 CTGAAAGGCTCCAGCAGCAGCTTCCACGTGGATCTGAGGACTTCCAGCAGCCTAGTGCT 498
QY 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
DB 499 CCTGGCTTTATTTGTGCAGAGTGGGAAAGACAGGCACCTAAGTCTGTGCAGACTTTACCG 558
QY 141 ProGluAspPheCysPheTrpGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
DB 559 CCAGAGGACTTCTGTCTTATCAAGGCTCTTTGGCATCACACAGAACTCTCTCACTGGCC 618
QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTrpPheLeu 180
DB 619 CTTTCACTGCGCAAGGCTTGTTCAGGCATGATACGACAGAGAGGCGAGTACTTCTTA 678
QY 181 ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro 200
DB 679 AGGCCACTCTCTCACCTCTCATGAAACTCGGCAGAGCTGCCCAAGCGCAGCTCGCCA 738
QY 201 SerHisValLeuTrpLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu 220
DB 739 TCCCACTGACTGTACAAGAGATCCACAGAGCCCCATGCTCTCTGGGGCGGCGAGTGTCTG 798
QY 221 ValThrSerArgTrpTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu 240
DB 799 GTGACCTCAGAGCATGGAGCTGGCAGCATCAACCTCTGCAGCAGCGACCTTCCGCTG 858
QY 241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTrpMetProGlnProPro 260
DB 859 GGACTGCCACAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCCAGCCTCC 918
QY 261 LysGluAspPheLeuLeuProAspGluTrpLysSerCysLeuArgHisLysArgSer 280
DB 919 AAGGAAGACCTCTTCTATCTTGCAGATGAGTAAAGTCTTGTCTTACGGCATAGCGCTCT 978

QY 281 LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValValAspLys 300
DB 979 CTTCTGAGTCCCATAGAAATGAAGACTGAACGTGGAGACCTTGGTGGTGGTGCACAAA 1038
QY 301 LysMetMetGlnAsnHisGlyHisGluAlaLeuThrTrpValLeuThrLeuLeuAsn 320
DB 1039 AAGATGATGCAAAACCATGGCCATGAAATATCACACCTACGTCTCACGATCTCAAC 1098
QY 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAlaIleVal 340
DB 1099 ATGGTATCTGCTTATTCAGAGATGGACATAGAGGAAACATCAACATTCGATTTGTA 1158
QY 341 GlyLeuLeuLeuGluAspGluGlnProGlyLeuValIleSerHisHisAlaAspHis 360
DB 1159 GGTCTGATTTCTTAGAAGATGAACAGCGCGCTGGTGTATAGTCACCCAGCAGACCAC 1218
QY 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
DB 1219 ACCTTAAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCA 1278
QY 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
DB 1279 GACCAAGCCATCTTACTGCTGCTGTGATATATGTTCTCTGGAAGATGAGCCCTGTGAC 1338
QY 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTrpArgSerCysThrIleAsn 420
DB 1339 ACTTTGGGATTTGCCACCCATTAAGTGAATGTGTAGTAAATATCGCAGCTGCACGATTA 1398
QY 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
DB 1399 GAAGATACAGGTCTTGGAGCTTCCCTTCCCATTTGCCATGAGTCTGGAACAACTTTGGC 1458
QY 441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
DB 1459 ATGATTCATGATGGAAGAGGAAACATGTGTAAGTTCGAGGCGCAACATCATGTCCCT 1518
QY 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTrpLeuHis 480
DB 1519 ACATTTGGCAGCAGCAATGGAGTCTTCTCTGGTCCACCTGCAGCCGCCAGTATCTAC 1578
QY 481 LysPheLeuArgSerValLysMetProAlaLeu 491
DB 1579 AAATTTCTAAGCACCGCTCAAGCTATCTGCCTTG 1612
RESULT 10
ACC00343
ID ACC00343 standard; DNA; 3227 BP.
XX ACC00343;
XX 30-JUN-2003 (first entry)
XX Human ADAMTS-16 coding sequence.
XX Human; cytostatic; antiarthritic; analgesic; antiinflammatory; leukaemia;
KW immunosuppressive; ADAMTS; disintegrin; metalloprotease; thrombospondin;
KW ADAMTS-15; ADAMTS-16; ADAMTS-17; ADAMTS-18; ADAMTS-19; wound healing;
KW extracellular matrix; cancer; metastasis; arthritis; pain; coagulation;
KW immune disorder; inflammation; immune response; angiogenesis; rheumatism;
KW embryo implantation; foetal development; nerve degeneration; gene; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 1..3219
XX /*tag= a
XX /product= "ADAMTS-16"
XX WO2003027282-A1.
XX 03-APR-2003.
XX

PF 24-SEP-2002; 2002WO-JF009771.
 XX
 PR 24-SEP-2001; 2001ES-00002165.
 PR 24-SEP-2001; 2001ES-00002166.
 PR 24-SEP-2001; 2001ES-00002167.
 PR 25-SEP-2001; 2001ES-00002192.
 PR 25-SEP-2001; 2001ES-00002193.
 XX (DAII-) DAIICHI FINE CHEM CO LTD.
 PA (UYOV-) UNIV OVIEDO.
 XX
 XX Cal Miguel S, Obaya Gonzalez AJ, Llamazares Prada M;
 PI Garabaya Fernandez C, Lopez-Otin C;
 XX
 DR MPI; 2003-313640/30.
 DR P-PSDB; ABR40093.
 XX
 XX ADAMTS family zinc metalloproteases for treatment and investigation of
 PT metalloprotease related disorders including cancer and arthritis.
 XX
 PS Claim 3; Page 118-122; 169pp; Japanese.
 XX
 CC The present invention relates to novel human ADAMTS (A disintegrin and
 CC metalloprotease domain-chondroponin domain) family proteins: ADAMTS-15,
 CC ADAMTS-16, ADAMTS-17, ADAMTS-18 and ADAMTS-19 (ACC00342-ACC00346 and
 CC ABR40092-ABR40096). The ADAMTS proteins and their coding sequences are
 CC useful for the treatment, prevention, diagnosis and investigation of
 CC disorders associated with the extracellular matrix, such as cancer,
 CC cancer metastasis, leukaemia, arthritis, pain, immune disorders and
 CC inflammation; investigation of related processes such as immune response,
 CC angiogenesis, coagulation, wound healing, embryo implantation, foetal
 CC development, cancer proliferation and infiltration, metastasis,
 CC rheumatism and nerve degeneration
 XX
 SQ Sequence 3227 BP; 776 A; 871 C; 898 G; 682 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,27e-173 Length: 3227
 Score: 2584.00 Matches: 484
 Percent Similarity: 99.19% Conservative: 4
 Best Local Similarity: 98.37% Mismatches: 3
 Query Match: 98.33% Indels: 1
 DB: 10 Gaps: 0
 US-10-804-457-2 (1-491) x ACC00343 (1-3227)
 QY 1 MetLysProArgAlaArgGlyTyrArgGlyLeuAlaLeuMetLeuLeuAlaGln 20
 DB 1 ATGAAGCCCGCGCGCGGATGCGGGCTTGGCGGCTGTGGATGCTGCTGGCGCAG 60
 QY 21 ValAlaGluGlnAlaProAlaCysAlaMetClyProAlaAlaAlaProGlySerPro 40
 DB 61 GTGGCCGAGCAGGACCTTGGCGGCGCATGGGACCCGCGCGCGCGCTGGGAGCCG 120
 QY 41 SerValProArgProProProProAlaGluArgProGlyTyrMetGluLysGlyGluTyr 60
 DB 121 AGCGTCCGCGCTCTCTCCACCCGCGAGCGCGCGCGGCTGGATGGAAAGGCGGATAT 180
 QY 61 AspLeuValSerAlaTyrGluAlaAspHisArgGlyAspTyrValSerHisGluIleMet 80
 DB 181 GACCTGTGTCTGTGCTACGAGGTGACACAGAGGCGGATTAACGTGTCCCATGAATCATG 240
 QY 81 HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg 100
 DB 241 CACCATCAGCGCGGAGAGGAGGAGGCGGCTGTCCGAGGTGAGTCTCTTACCTTCG 300
 QY 101 LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerLeuValAla 120
 DB 301 CTGAAGGCGCGCGAGGACGATCTCCATGATGATGATGATGATGATGATGATGATGAT 360
 QY 121 ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro 140
 DB 361 CCTGGCTTATTTGTGAGAGCTTTGGGAAAGACAGGCGCATTAAGTGTGTGCGAGAT 420

QY 141 ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla 160
 DB 421 CCAGAGGAGCTTCTGTTTCTATCAAGGCTCTTTCGATCAGAGAACTCTCTAGTGGCC 480
 QY 161 LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu 180
 DB 481 CTTTCAACCTGCCAAGGCTTGTGAGGATGATAGCAACAGAGAGGAGATTAATCTCTTA 540
 QY 181 ArgProLeuProSerHisLeuSerTyrPheLeuGlyValGalaAlaGlnGlySerSerPro 200
 DB 541 AGGCCACTTCTTCAACCTCTCATGGAACTCCGAGAGCTGCCAAGGAGCTCGCCA 600
 QY 201 SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyValAserGluValLeu 220
 DB 601 TCCACAGCTACTGTATCAAGAGATCCACAGAGCCCCATGCTCTCTGGGCCAGTGAGGCTCTG 660
 QY 221 ValThrSerArgThrTyrGluLeuAlaHisGlnProLeuHisSerSerArgLeuArgLeu 240
 DB 661 GTGACCTCAAGGACATGGAGCTGGCACATCAACCCCTGCACAGCAGCAGCTTCGCGCTG 720
 QY 241 GlyLeuProGlnLysGlnHisPheCysGlyValArgLysLysTyrMetProGlnProPro 260
 DB 721 GGACTGCCACAAAAGCAGCATTCTGTGGAGAGCGAGCAAGAAATACATGCCAGCCCTCCC 780
 QY 261 LysGluAspLeuPheHelleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
 DB 781 AAGGAGACCTCTTCATCTTCCAGATGAGTATAAGTCTTGTCTTACGGCATAAGCGCTCT 840
 QY 281 LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValAlaAspLys 300
 DB 841 CTTCTGAGGTCCCATAGAAATGAAGAACTGAAGCTGGAGACCTTGGTGGTGGTGGACAAA 900
 QY 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrThrTyrValLeuThrIleLeuAsn 320
 DB 901 AAGATGATGCAAAACCATGGCCATGAAATATACCCACCTACGTGCTCAGCATCTCAAC 960
 QY 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyCysAsnIleAsnIleAlaIleVal 340
 DB 961 ATGGTATCTGCTTATTAAGATGAAGCAAAATAGGAGGAAACATCAACATGCAATTTGA 1020
 QY 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis 360
 DB 1021 GGTCTGATTTCTTAGAAGATGAACAGCAGGAGCTGTGATTAAGTCAACCCAGCAGACC 1080
 QY 361 ThrLeuSerSerPheCysGlnTyrGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
 DB 1081 ACCTTAAGTAGCTTCTGCGAGTGGCAGTCTGGATTAATGGGAAAGATGGGACTCGTCAT 1140
 QY 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTyrLysAsnGluProCysAsp 400
 DB 1141 GACCAAGCCATCTTACTGACTGTCTGGATATATGTTCTCTGGAAGATGAGCCCTGTGAC 1200
 QY 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
 DB 1201 ACTTTGGGATTTGGCAACCCATAGTGAATGTAGTAAATATCGCAGCTGCAGATTAAT 1260
 QY 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerClyHisAsnPheGly 440
 DB 1261 GAAGATACAGGCTTGGAGCTGGCTTCCATTTGCCATGAGCTGGAGCACAACTTTGGC 1320
 QY 441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
 DB 1321 ATGATTCATGATGGAGAGGGAACATGTGTAAAGTCCGAGGGCAACATCATGTCTCCCT 1380
 QY 461 ThrLeuAlaGlyArgAsnGlyValPheSerTyrPheSerProCysSerArgGlnThrLeuHis 480
 DB 1381 ACATGGCAGGACGCAATGGAGTCTTCTCTGGTCAACCTCGAGCCGCGAGTATCTACAC 1440
 QY 481 LysPheLeuArgSer-ValLysMetProAlaLeu 491
 DB 1441 AAATTTCTAAGCAGCGCTCAAGCTATCTGCGCTTG 1474

RESULT 11

AAS97174
ID AAS97174 standard; cDNA; 3675 BP.

AC AC
AAS97174;

DT 26-FEB-2002 (first entry)

XX Human metalloprotease partial DNA sequence #3.

XX Human; protease; PCR primer; cytostatic; immunomodulator; cardiant;
KW vasotropic; antimigraine; analgesic; endocrine; nootropic; tranquiliser;
KW hypertensive; hypotensive; neuroleptic; neuroprotective; anabolic;
KW anorectic; antiinflammatory; aspartyl protease; cysteine protease;
KW metalloprotease; serine protease; cancer; haematopoietic; breast; colon;
KW lung; prostrate; cervical; brain; ovarian; bladder; kidney; pain;
KW immune-related disease; cardiovascular disease; neuronal disease;
KW migraine; sexual dysfunction; mood disorder; attention disorder;
KW cognition disorder; hypotension; hypertension; psychotic disorder;
KW dyskinesia; metabolic disorder; inflammatory disorder; ss.

XX Homo sapiens.

XX WO200183782-A2.

XX 08-NOV-2001.

XX 04-MAY-2001; 2001WO-US014431.

XX 04-MAY-2000; 2000US-0201879P.

XX (SUGEN-) SUGEN INC.

XX Plowman GD, Whyte D, Sudarsanam S, Manning G, Caenepeel S;
PI Payne V;

XX WPI; 2002-041502/05.

XX P-PSDB; AAU72891.

XX Novel protease polypeptide useful for screening for substances that may
PT be used to treat, e.g., cancers, immune-related diseases, cardiovascular
PT disease, migraine, pain, psychotic and inflammatory disorders.

XX Claim 30; Fig 1H; 232pp; English.

XX The invention relates to an isolated, enriched, or purified protease
CC polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to
CC screen for substances (S) that may modulate its activity. Administering S
CC (which modulates protease activity in vitro) may be used to treat a
CC disease or disorder selected from cancers (e.g., of tissues, of blood or
CC haematopoietic origin, of the breast, colon, lung, prostate, cervical,
CC brain, ovarian, bladder or kidney), immune-related diseases and
CC disorders, cardiovascular disease, brain or neuronal-associated diseases
CC (e.g., central or peripheral nervous system diseases, migraine, pain,
CC sexual dysfunction, mood disorders, attention disorders, cognition
CC disorders, hypotension, hypertension, psychotic disorders, neurological
CC disorders and dyskinesias), metabolic disorders and inflammatory
CC disorders. (II) may also be useful as a diagnostic tool for a disease or
CC disorder such as those above. AAS97159-AAS97195 represent human protease
CC coding sequences and primers of the invention

XX Sequence 3675 BP; 874 A; 1009 C; 1019 G; 773 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,67e-173	Length:	3675
Score:	2584.00	Matches:	484
Percent Similarity:	99.19%	Conservative:	4
Best Local Similarity:	98.37%	Mismatches:	3
Query Match:	98.33%	Indels:	1
DB:	6	Gaps:	0

US-10-804-457-2 (1-491) x AAS97174 (1-3675)

QY	1	MetLysProArgAlaArgGlyTTPArgGlyLeuAlaAlaLeuTTPMetLeuLeuAlaGln	20
DB	1	ATGAAGCCCGCGCGCGGATGGCGGGCTTGGCGGCGCTGTGGATGTGTGGCGAG	60
QY	21	ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerPro	40
DB	61	GTGGCCGAGCGGACCTTGGTGGCCATGGAGCCCGGAGCGGCGCTGGGAGCCCG	120
QY	41	SerValProArgProProProAlaGluArgProGlyTTPMetGluLysGlyGluTyr	60
DB	121	AGCGTCCCGGCTCTCTCCACCGCGAGCGCGGCTGGATGGAAAGCGCGAATAT	180
QY	61	AspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluLeuMet	80
DB	181	GACCTGGTCTCTGGCTACGAGGTTCACACAGGGCGATTACGTGTCCCATGAATCATG	240
QY	81	HisHisGlnArgArgArgAlaValSerGluValGluSerLeuHisLeuArg	100
DB	241	CACCATCAGCGCGGAGAGAGAGTGGCGGTCCGAGGTGGAGTGTCTTTCACCTTCGG	300
QY	101	LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerLeuValAla	120
DB	301	CTGAAGGCCCGCAGCAGCTTCCACATGGATCTGAGGACTTCCAGCAGCTAGTGGCT	360
QY	121	ProGlyPheLeValGlnThrLeuGlyThrGlyThrLysSerValGlnThrLeuPro	140
DB	361	CCTGGCTTTATTGTGCAGACGTTGGGAAAGACAGGCACCTAAGTCTGTGCAGACTTTACCG	420
QY	141	ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerValAla	160
DB	421	CCAGAGGACTTCTGTTCATCAGGCTTTTGGCCATCACAGAAACTCTCTCAGTGGCC	480
QY	161	LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu	180
DB	481	CTTCAACTGCCAAGGCTTGTACGCATGTACGACAGAGAGGAGGAGTACTTCTCTA	540
QY	181	ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro	200
DB	541	AGGCCACTTCTCTCACACTCTCATGGAAACTCGGCAGAGCTGCCAGGAGCTGCCA	600
QY	201	SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu	220
DB	601	TCCCACTACTGTACAAGATCCACAGGCCCATGCTCTCTGGGGCCAGTGAGGTCTTG	660
QY	221	ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu	240
DB	661	GTGACTCTCAGGACATGGAGCTGGCACATCAACCCCTGCACAGCAGGACCTTCGCTG	720
QY	241	GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro	260
DB	721	GGACTGCCCAAAAGCAGCATTTCTGTGAAGACGCAAGAAATACATGCCCGAGCTCCC	780
QY	261	LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer	280
DB	781	AAGGAAGACCTCTTCATCTTGCAGATGAGTATAAGTCTTGTCTTACGGCATTAAGGCTCT	840
QY	281	LeuLeuArgSerHisArgAsnGluLeuLeuValGluThrLeuValValValAspLys	300
DB	841	CTTCTGAGGTCCCATAGAAATGAGAACTGAACGTGGAGACCTTGGTGGTGCACAAA	900
QY	301	LysMetMetGlnAsnHisGlyHisGluAsnIleThrThrTyrValLeuThrIleLeuAsn	320
DB	901	AAGATGATGCAAAACCATGGCCATGAAAATATCACCACTACGTCTCACGATACTCAAC	960
QY	321	MetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAlaIleVal	340
DB	961	ATGGTATCTGTCTTATTCAAGATGGAACAATAGGAGGAAACATCAACATTGCAATTGTA	1020
QY	341	GlyLeuLeuLeuLeuGluAspGlnProGlyLeuValIleSerHisHisAlaAspHis	360
DB	1021	GGTCTGATTCTCTAGAGATGAACAGCAGGACTGGTGATAGTCAACACCGACACAC	1080
QY	361	ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis	380

Db	1081	ACCTTAAGTAGCTTCTCCAGTGGCAGCTCGGATTGATGGGAAGATGGAGCTCGTCAT	1140
Qy	381	AspHisAlaIleLeuLeuThrGlyVLeuAspIleCysSerTrpIysAenGluProCysAsp	400
Db	1141	GACCAGCCCATCTTACTGACTGGTCTGGATATATGTTCTTGGAAAGAAATGAGCCCTGTGAC	1200
Qy	401	ThrLeuGlyPheAlaProIleSerGlyMetCysSerIysTrpArgSerCysThrIleAsn	420
Db	1201	ACTTTGGGATTTGCACCCATAGTGGAAATGTAGTAAATATCGAGTCGACGATTAAT	1260
Qy	421	GluAspThrGlyVLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAenPheGly	440
Db	1261	GAAGATACAGGCTCTTGGACTGGGCTTCACATTGCCATGAGTCTGGACACAACCTTTGGC	1320
Qy	441	MetIleHisAspGlyGluGlyAenMetCysLysLysSerGluGlyAenIleMetSerPro	460
Db	1321	ATGATTTCATGATGGAGAAGGGAAACATGTGTAAANAAGTCGAGGGCAACATCATGTCCCT	1380
Qy	461	ThrLeuAlaGlyArgAenGlyValPheSerTrpSerProCysSerArgGlnTrpLeuHis	480
Db	1381	ACATTGGCAGGACGCAATGGAGTCTTCTCTGTGTACCCCTGCGAGCCGCCAGTATCTACAC	1440
Qy	481	LysPheLeuArgSer-ValLysMetProAlaLeu	491
Db	1441	AAATTTCTAAGCACCGCTCAAGCTATCTGCCTTG	1474

RESULT 12	
ACF57047	
ID	ACF57047 standard; cDNA, 3675 BP.
XX	
AC	ACF57047;
XX	
DT	09-OCT-2003 (first entry)
XX	
DE	Human ADAMTS16 encoding cDNA SEQ ID NO:1.
XX	
KW	Human; aggrecanase; enzyme; ADAMTS; ADAMTS16; osteopathic; antiarthritic;
KW	antiinflammatory; gene therapy; osteoarthritis; gene; ss.
XX	
OS	Homo sapiens.

PN	WO2003057842-A2.
XX	
XX	17-JUL-2003.
XX	
XX	30-DEC-2002; 2002WO-US041730.
XX	
XX	31-DEC-2001; 2001US-0344895P.
XX	
XX	(AMHP) WYETH.
XX	
XX	Agostino MJ, Diblasio E;
XX	
XX	WPI; 2003-577519/54.
XX	P-PSDB; ABR83432.
XX	

comprises: (a) providing an aggrecanase protein (fragment); (b) combining the aggrecanase with a potential inhibitor; and (c) evaluating whether the potential inhibitor inhibits aggrecanase activity, where the aggrecanase protein is used in a three dimensional structural analysis prior to combining with the potential inhibitor, or is used in a computer aided drug design prior to combining with the potential inhibitor. A composition comprising an antibody that binds to a purified aggrecanase protein can be used for inhibiting aggrecanase activity in a mammal, for treatment of aggrecanase-associated conditions, e.g. osteoarthritis

Sequence 3675 BP; 874 A; 1009 C; 1019 G; 773 T; 0 U; 0 Other;

```
QY 261 LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db 781 AAGGAAGACCTCTTCACTTCGTCAGATGAGTATAAGTCTTGCTTACGGCATAAGCGCTCT 840
QY 281 LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValAlaAspLys 300
Db 841 CTTCTGAGTCCCATAGAATGAAGAACTGAACGTGGAGACCTTGTTGGTGTGCGCAAA 900
QY 301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrTyrValLeuThrIleLeuAsn 320
Db 901 AAGATGATGCMAAACCATGGCCATGAATATATACCACTACCTGCTCAGTACTCAAC 960
QY 321 MetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAlaIleVal 340
Db 961 ATGGTATCTGCTTATTCAAGATGGAACAATAGGAGGAACATCAACATTCGAATTGTA 1020
QY 341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis 360
Db 1021 GGTCTGATTCTTTAGAGATGAACAGCCAGGACTGGTGTATAGTCACCACGAGACCAC 1080
QY 361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db 1081 ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTGCTCAT 1140
QY 381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db 1141 GACCAGCCATCTTACTGACTGCTCTGGATATATATGTTCTGGAAGAATGAGCCCTGTGAC 1200
QY 401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
Db 1201 ACTTTGGGATTTGCAACCCATAAGTGAATGTGTAGTAAATATATGCGACGCTGCACGATTAA 1260
QY 421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db 1261 GAAGATACAGGCTTGGATGGCCCTTACCATTGCCCATGAGTCTGGACACAACTTTGGC 1320
QY 441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
Db 1321 ATGATTCTATGAGAGAGGAACATGTTGTAAAAAGTCCGAGGCAACATCATGTCCTCT 1380
QY 461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
Db 1381 ACATTGGCAGGACGCAATGAGTCTTCTCTGTCACCTGCGAGCGCCAGTATCTATACAC 1440
QY 481 LysPheLeuArgSerValLysMetProAlaLeu 491
Db 1441 AAATTTCATGACACCGCTCAAGCTATCTGCTTG 1474
```

RESULT 13

```
ABK49822
ID ABK49822 standard; cDNA; 5610 BP.
AC AC ABK49822;
XX XX
DT DT
DE DE
XX XX
XX XX
KW Human; ss; gene; ADAMTS; cytostatic; antidiabetic; antirheumatic; SNP;
KW antiarthritic; antiulcer; vulnery; neovascularisation; angioma;
KW diabetic omentopathy; chronic rheumatoid arthritis; gene therapy;
KW refractory skin ulcer; gastric ulcer; post-operative healing failure;
KW reprolysin-type 2N-metalloprotease domain; disintegrin-like domain; TSPl;
KW thrombospondin type 1 domain; sexual cycle; tumour; 5P-syndrome deletion;
KW chromosome 5p15.2-15.3; Cri-du-chat syndrome;
XX single nucleotide polymorphism.
XX Homo sapiens.
XX XX
FH Key Location/Qualifiers
FT CDS 770..4444
FT /*tag= a
```

```
FT variation /product= "ADAMTS"
FT replace(821,T)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism"
FT variation /standard_name= "Single nucleotide polymorphism"
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism"
FT variation /standard_name= "Single nucleotide polymorphism"
FT /*tag= d
FT /standard_name= "Single nucleotide polymorphism"
FT variation /standard_name= "Single nucleotide polymorphism"
FT /*tag= e
FT /standard_name= "Single nucleotide polymorphism"
FT variation /standard_name= "Single nucleotide polymorphism"
FT /*tag= f
FT /standard_name= "Single nucleotide polymorphism"
XX WO200231163-A1.
PN 18-APR-2002.
XX 11-OCT-2001; 2001WO-JP008913.
XX 11-OCT-2000; 2000JP-00311309.
PR 02-APR-2001; 2001JP-00102905.
XX (KAZU-) KAZUSA DNA RES INST FOUND.
XX (MITS-) MITSUBISHI PHARMA CORP.
XX Ohara O, Nagase T, Nomura N, Yano K, Murakami K, Yasuda S;
XX Kanzaki K;
XX WPI; 2002-372277/40.
XX P-PSDB; AAU79497.
XX Human brain-originated ADAMTS family polypeptide and encoded gene,
XX applicable in diagnosis and screening compounds for drug compositions in
XX treating diseases due to e.g. neovascularization.
XX Claim 4; Page 135-145; 172pp; Japanese.
XX The invention relates to a polypeptide belonging to the ADAMTS family is
XX selected from sequences appearing as AAU79496, AAU79497 and AAU79499, a
XX protein that contains the polypeptide, a protein having not less than 50%
XX homology with the amino acid sequence of the polypeptides or a
XX polypeptide modified from any of the polypeptides but with some amino
XX acids deleted, substituted, added or inserted. Also included are the
XX polynucleotides encoding the polypeptides (or their complementary strands
XX or variants), a recombinant vector containing any of the polynucleotides,
XX a transformant which is transformed with the recombinant vector,
XX producing the polypeptide, protein or peptide by culturing the
XX transformant, an antibody that can recognize the polypeptide, protein or
XX peptide and screening compounds to promote or inhibit activity of the
XX polypeptide or protein, or to promote or inhibit expression of the
XX polynucleotide by using the polypeptide, protein, peptide, the
XX polynucleotide, vector, transformant or/and antibody, particularly in the
XX presence of a test compound for contact before evaluating the activity by
XX measuring signal changes. The polypeptide and encoded gene are applicable
XX in diagnosis and screening compounds for drug compositions in treating
XX diseases due to neovascularisation, diabetic omentopathy, chronic
XX rheumatoid arthritis, angioma, refractory skin and gastric ulcers and
XX post-operative healing failure, including gene therapy. The gene encoding
XX such polypeptide has conserved reprolysin-type 2N- metalloprotease
XX domain, disintegrin-like domain and TSPl (thrombospondin type 1) domain.
XX Its encoded protein is characterised by high expression in ovaries,
XX changes in expression dose depending on the sexual cycle, a decrease in
XX tumour cell and location of the gene on the 5P-syndrome deletion site on
XX chromosome 5p15.2-15.3 (associated with Cri-du-chat syndrome). The
XX present sequence encodes a ADAMTS protein of the invention
XX SQ Sequence 5610 BP; 1233 A; 1638 C; 1537 G; 1202 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.:	4.5e-173	Length:	5610
Score:	2584.00	Matches:	484
Percent Similarity:	99.19%	Conservative:	4
Best Local Similarity:	98.37%	Mismatches:	3
Query Match:	98.33%	Indels:	1
DB:	6	Gaps:	0

US-10-804-457-2 (1-491) x ABK49822 (1-5610)			
QY	1	MetLysProArgAlaArgGlyTrpArgGlyLeuAlaAlaLeuTrpMetLeuLeuAlaGln	20
DB	770	ATGAGCCCGCGCGCGGATGGCGGCGCTGGCGGCTGGGATGCTGGCGCGAG	829
QY	21	ValAlaGluGlnAlaProAlaCysAlaMetGlyProAlaAlaAlaAlaProGlySerPro	40
DB	830	GTGGCCGAGCAGCACCTGCGTGGCCCATGGGACCCCGCAGCGCGCTGGGAGCCCG	889
QY	41	SerValProArgProProProAlaGluArgProGlyTrpMetGluLysGlyLys	60
DB	890	AGCGTCCCGGCTCTCTCCACCCGCGAGCGCGCGGCTGGATGGAAGGCGCAATAT	949
QY	61	AspLeuValSerAlaTyrGluValAlaPheHisArgGlyAspTyrValSerHisGluLeuMet	80
DB	950	GACCTGGTCTTGGCTTACGAGGTTGACACAGGGCGGATTAAGTGTCCCATGAAATCATG	1009
QY	81	HisHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeuArg	100
DB	1010	CACCATCAGCGCGGAGAGAGCAGTGGCGGCTGTCGAGGTTGAGTCTCTTCACCTTCGG	1069
QY	101	LeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuValAla	120
DB	1070	CTGAAGGCGCCAGGACGACACTTCCACATGGATCTGAGGACTTCCAGCAGCGCTAGTGCT	1129
QY	121	ProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuPro	140
DB	1130	CCTGGCTTATTGTGCAGACGTTGGGAAAGACAGGACACTAAGTCTGTGCAGACTTTACG	1189
QY	141	ProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerValAla	160
DB	1190	CCAGAGGACTTCTGTGTTCTATCAGGCTCTTTGGCATCACACAGAAACTCTCTCAGTGCC	1249
QY	161	LeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeu	180
DB	1250	CTTTCAACTGCCAGGCTTGTGAGGCAATGATACAGAGAGGCGAGATTACTTCCCTA	1309
QY	181	ArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerPro	200
DB	1310	AGGCCACTTCTCTCACACTCTCATGGAAACTCGGAGAGCTGCCAAGCGAGCTCGCCA	1369
QY	201	SerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluValLeu	220
DB	1370	TCCACAGTACTGTACAGAGATCCACAGAGCCCATGCTCTGGGGGCGAGTGAGGTCCTG	1429
QY	221	ValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeu	240
DB	1430	GTGACCTCAAGGACATGGAGCTGGCACAATCAACCTTCACAGCAGCGACCTTCGGCTG	1489
QY	241	GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro	260
DB	1490	GGACTGCCACAAAAGCAGCACTTTCTGTGGAAGACGCAAGAAATACATGCCCCAGCTCCC	1549
QY	261	LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer	280
DB	1550	AAGGAGACCTTCTTCATCTTGCCAGATGAGTATAAGTCTTGGCTTACGGCATAAGCGCTCT	1609
QY	281	LeuLeuArgSerHisArgAsnGluLeuAsnValGluThrLeuValValValAspLys	300
DB	1610	CTTCTGAGGTCCCATAGAAATGAAGAACTGAACCTGGAGACCTTGGTGGTGGTCGACAA	1669
QY	301	LysMetMetGlnAnHisGlyHisGluAsnIleThrThrTyrValLeuThrIleLeuAsn	320
DB	1670	AAGATGATGCAAAACCATGGCCATGAAATATCACCACTACGTGCTCACGATACTCAAC	1729

QY	321	MetValSerAlaLeuPheLysAspGlyThrIleGlyLysAsnIleAsnIleAlaIleVal	340
DB	1730	ATGGTATCTGCTTTATTCAAAGATGGAACAATAGGAGGAAACATCAACATTGCAATTGTA	1789
QY	341	GlyLeuIleLeuLeuLeuGluAspGluGlnProGlyLeuValIleSerHisAlaAspHis	360
DB	1790	GGTCTGATCTTCTAGAAGATGAACAGCCAGGACTGGTGATAAGTCAACGACGACAC	1849
QY	361	ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis	380
DB	1850	ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGGAAGATGGGACTGCTCAT	1909
QY	381	AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp	400
DB	1910	GACCACGCCCATCTTACTGACTGGTCTGGATATATGTTCTCTGGAAGATGAGCCCTGTGAC	1969
QY	401	ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn	420
DB	1970	ACTTTGGGATTTGCACCATTAAGTGGAAATGTGTAGTAAATATCCGAGCTGCACGATTAA	2029
QY	421	GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly	440
DB	2030	GAAGATACAGTCTTGGACTGGCTTCCACATTGCCCATGAGTCTGGACACAACTTTGGC	2089
QY	441	MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro	460
DB	2090	ATGATTATGATGGAAGGGAACATGTGTAAAGTCCGAGGGAACATCATGTCCCT	2149
QY	461	ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis	480
DB	2150	ACATTGCGAGGACGCAATGAGTCTTCTCTGGTCCACCTTCAGCCGCCAGTATCTACAC	2209
QY	481	LysPheLeuArgSerValLysMetProAlaLeu	491
DB	2210	AAATTTCTTAAGCACCGCTCAAGCTATCTGCTTGG	2243

RESULT 14	
ABK49825	
ID	ABK49825 standard; DNA, 8435 BP.
XX	
AC	ABK49825;
XX	
DT	15-JUL-2002 (first entry)
XX	
DE	Plasmid pFastBac1-HT-PJ01256.
XX	
KW	ds; ADAMTS; cytostatic; antidiabetic; antirheumatic; cyclic;
KW	antiarthritic; antiulcer; vulnery; neovascularisation; angioma;
KW	diabetic omentopathy; chronic rheumatoid arthritis; gene therapy;
KW	refractory skin ulcer; gastric ulcer; post-operative healing failure;
KW	reprolysin-type 2N-metalloprotease domain; disintegrin-like domain; TSPI;
KW	thrombospondin type 1 domain; sexual cycle; tumour; 5P-syndrome deletion;
KW	chromosome 5p15.2-15.3; Cri-du-chat syndrome; circular;
XX	pFastBac1-HT-PJ01256.
XX	
OS	Homo sapiens.
OS	Escherichia coli.
OS	unidentified baculovirus.
OS	Rhesus macaque polyoma virus.
OS	Transposon Tn7.
OS	Synthetic.
PN	WO200231163-A1.
XX	
PD	18-APR-2002.
XX	
PF	11-OCT-2001; 2001WO-JP0089913.
XX	
PR	11-OCT-2000; 2000JP-00311309.
PR	02-APR-2001; 2001JP-00102905.
XX	
PA	(KAZU-) KAZUSA DNA RES INST FOUND.
PA	(MITS-) MITSUBISHI PHARMA CORP.


```
QY      241 GlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetProGlnProPro 260
Db      4842 GGACTGCCACAAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCGCTCCC 4901
QY      261 LysGluAspLeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSer 280
Db      4902 AAGGAAGACCTCTTCATCTTGCAGATGAGTATAGTCTTGTACGGCATTAAGGCTCT 4961
QY      281 LeuLeuArgSerHisArgAsnGluGluLeuAsnValGluThrLeuValValAspLys 300
Db      4962 CTTCTGAGGTCCCATAGAAATGAAGAACTGAACGTGGAGACCTTGTGTGTGTCGACAAA 5021
QY      301 LysMetMetGlnAsnHisGlyHisGluAsnIleThrThrTyrValLeuThrIleLeuAsn 320
Db      5022 AAGATGATGCAAAACCATGGCCATGAAATATATCACCACCTAGCTGTCTACGATACTCAAC 5081
QY      321 MetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAsnIleAlaIleVal 340
Db      5082 ATGGTATCTGCTTTATTCAAGATGGAACAATAGGAGGAACATCAACATTGCAATTGTA 5141
QY      341 GlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisHisAlaAspHis 360
Db      5142 GGTCTGATTTCTTAGAAGATGAACAGCCAGGACTGGTGTATAGTCACCCAGCAGCCAC 5201
QY      361 ThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHis 380
Db      5202 ACCTTAAGTAGCTTCTCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCAT 5261
QY      381 AspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAsp 400
Db      5262 GACCCAGCCATCTTACTGACTGGTCTGGATATATGTCTCTGGAAGAAATGAGCCCTGTGAC 5321
QY      401 ThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsn 420
Db      5322 ACTTTGGGATTTGCACCCATAGTGGAAATGTGTAGTAATAATCGCAGCTGCACGATTAT 5381
QY      421 GluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGly 440
Db      5382 GAAGATACAGGCTCTTGGACTGGCCCTTCCACCATGGCCATGAGTCTGGACACAACCTTTGGC 5441
QY      441 MetIleHisAspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerPro 460
Db      5442 ATGATTTCATGTGAGAGGGAAGCAACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCT 5501
QY      461 ThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHis 480
Db      5502 ACATTGGCAGGACCAATGGAGTCTTCTCTGCTACCCCTGCAGCCGCGAGTATCTACAC 5561
QY      481 LysPheLeuArgSer-ValLysMetProAlaLeu 491
Db      5562 AAATTTCTAAGCACCGCTCAAGCTATCTGCCCTTG 5595
```

Search completed: June 14, 2005, 11:36:37
Job time : 709 secs

THIS PAGE BLANK (USPTO)


```

QY 365 PheCysGlnTrpClnSerGlyLeuMetGlyLysAspGlyThrArgHisAspHisAlaIle 384
Db 1310 TTCTGCCAGTGGCAATCCGGATTGATGGGGAAGATGGAACTCGTCATGACCACGCCATC 1369

QY 385 LeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAspThrLeuGlyPhe 404
Db 1370 TTAAGTCTGGTCTGACATATGTTCTCTGGAGAAATGAGCCCTGTGACACATTTGGGGTTT 1429

QY 405 AlaProIleSerGlyMetCysSerLysTrpArgSerCysThrIleAsnGluAspThrGly 424
Db 1430 GCACCCATAAGCGGATGTCAGTAAGTACCGCAGCTGCACACATCAATGAAGACTCAGGA 1489

QY 425 LeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGlyMetIleHisAsp 444
Db 1490 CTTGGACTGGCCCTTCCCATTCGCCATGAGTCTGGACACAACTTTGGCATGGTCCATGAT 1549

QY 445 GlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerProThrLeuAlaGly 464
Db 1550 GGAGAGGGAATATGTGCAAGAAATCTGAGGGCAATATATGTGCCCAACACTGGCAGGA 1609

QY 465 ArgAsnGlyValPheSerTrpSerProCysSerArgGlnTrpLeuHisLysPheLeuArg 484
Db 1610 CGCAATGGTGTCTTCTCTGGTCTTCTCGACGCCGTCACTACTGTCACAACTTTTAAGC 1669

QY 485 SerValLys 487
Db 1670 ACCGCCCAA 1678

RESULT 2
BM906555
LOCUS BM906555
DEFINITION AGENCOURT_6621721 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5722286
5', mRNA sequence.
ACCESSION BM906555
VERSION BM906555.1 GI:19356934
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
MAMMALIA; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 1122)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Cloned by: Agencourt Bioscience Corporation
Cloned distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M12708 row: b column: 15
High quality sequence start: 19
High quality sequence stop: 633.
Location/Qualifiers
1..1122
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5722286"
/lab_host="DH10B"
/clone_lib="NIH_MGC_125"
/notes="Organ: Ovary (pool of 3); Vector: pCMV-SPORT6;
Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(Score) site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics

```

FEATURES

source

tracking code 036."

```

ORIGIN
Alignment Scores:
Pred. No.: 1,94e-106 Length: 1122
Score: 1303.00 Matches: 275
Percent Similarity: 84.18% Conservative: 7
Best Local Similarity: 82.09% Mismatches: 43
Query Match: 49.58% Indels: 10
DB: 5 Gaps: 3
US-10-804-457-2 (1-491) x BM906555 (1-1122)

QY 73 AspTyrValSerHisGluIleMetHisHisGlnArgArgArgAlaValAlaValSer 92
Db 3 GATTACGTTGCCCTCAAAATCATGCACCATCAGCGCGGAGAGACAGTCGCCGTGTC 62

QY 93 GluValGluSerLeuHisLeuArgLeuLysGlySerArgHisAspPheHisValAspLeu 112
Db 63 GAGGTTGAGTCTCTTCCACTTCGGCTGAAGGCTCCAGGCACGACTTCCACGTGGATCTC 122

QY 113 ArgThrSerSerSerLeuValAlaProGlyPheIleValGlnThrLeu-GlyLysThrG 132
Db 123 GGGACTTCCAGCAGCTAGTGGCTCTCTGGCTTTATTGTGCAGACGTTGCGGAAGACAG 182

QY 132 YThrLysSerValGlnThrLeuProGluAspPheCysPheTyrGlnGlySerLeuAr 152
Db 183 CACTAAGTCTGTGCAGACTTTTACCGCCAGAGGACTTCTGTTCTATCAAGGCTCTTTGCG 242

QY 152 GSerHisArgAsnSerSerValAlaLeuSerThrCysGlnGlyLeuSerGlyMetIleAr 172
Db 243 ATCACACAGAACTCTCTAGTGGCCCTTTCACCTGCCAAGGCTTGTGAGGCATGATACG 302

QY 172 GThrGluGluAlaAspTyrPheLeuArgProLeuProSerHisLeuSerTrpLysLeuG 192
Db 303 ACAGAGAGGCGAGATTACTTCTTAAGGCCACTTCTCTCACACTCTCATGGAACCTCG 362

QY 192 YArgAlaGlnGlySerSerProSerHisValLeuTyrLysArgSerThrGluProHi 212
Db 363 CAGAGTGTGCCAAGCGCAGCTCGCCATCCACGCTACTGTACAGAGATCCACAGAGCCCA 422

QY 212 AlaProGlyAlaSerGluValLeuValThrSerArgThrTrpGluLeuAlaHisGlnPr 232
Db 423 TGCTCTGGGGCAGTGAGGCTCTGTGTGACCTCAAGGACATGGAGCTGGCAGTCAACC 482

QY 232 OleuHisSerSerAspLeuArgLeuGlyLeuProGlnLysGlnHisPheCysGlyArgAr 252
Db 483 CTGCACAGCAGCGACCTTGGCTGGGACTGCCAAGAGAGCAGCATTTCTGTGAAGACG 542

QY 252 GlyLysTyrMetProGlnProProLysGluAspLeuPheLeuLeuProAspGluTyrLy 272
Db 543 CAAGAAATACATGTCGCCAGCTCCCAAGGAAGACCTCTTTCATCTTCCAGATGAGTATAA 602

QY 272 SSerCysLeuArgHisLysArgSerLeuLeuArgSerHisArgAsnGluLeuAsnVa 292
Db 603 GTCTTGTCTTACGCATTAAGCGCTCTCTCTGAGGTGCCATAGAAATGAAGAACTGAACGT 662

QY 292 lGluThrLeuValValValAspLysLysMetMetGlnAsn-HisGlyHisGluAsnIle 312
Db 663 GGAGACCTTGGTGGGGTGCACAAAGATGATGCAAAACCCATGCCCATGAAATATCA 722

QY 312 hr-ThrTyrValLeuThrIleLeuAsnMetVal-SerAla-LeuPheLysAspGlyThrI 331
Db 723 CCCCCCTACGGGCTCACCATCTCAACTGTTAACTGCTGCTTTTATTATCAAGAGGGAACAA 782

QY 331 leGlyGlyAsnIleAsnIleAla---IleValGlyLeuLeuLeuGluAspGluGlnP 350
Db 783 TAGGGGGGAACCTTCAACTTGGCAATTGGACGGGTGATTCCTCTTAAAGATGAACAGC 842

QY 350 ro-GlyLeuValIleSerHisAlaAspHis---ThrLeuSerSerPheCysGlnTrp 368
Db 843 CACGAGCTGGGGGTAAGTTTCCCCCGGACCCCCCTTTAGATAGCTTCTGCG----- 896

```

Qy 369 GlnSerGlyLeuMetGlyLysaspGlyThrArgHisAAspHisAlaIleLeuLeuThrGly 388
 Db 897 CCAGGGGGAATTTGGAATTGATGGGGAAGGGAAGGAACTCTTCTGTGCAACGC 956
 Qy 389 LeuAspIleCysSerTrpLysAsnGluProCys 399
 Db 957 CCATCTTATCTACTGACTGCCCGGGAATAATCT 989

RESULT 3
 AK035797
 LOCUS
 DEFINITION
 Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:9630005B22 product:hypothetical Metalloprotease (ADAM type)/reprolysin (M12B) family containing protein, full insert sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AK035797 3399 bp mRNA linear HTC 03-APR-2004
 Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:9630005B22 product:hypothetical Metalloprotease (ADAM type)/reprolysin (M12B) family containing protein, full insert sequence.

AK035797.1 GI:26330919
 HTC; CAP trapper.
 Mus musculus (house mouse)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159

2 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076661

3 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection Nature 409, 695-690 (2001)

4 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
 6 (bases 1 to 3399)

5 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, K., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE
 JOURNAL

Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@cc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.jp/
 URL: http://fantom.gsc.riken.jp/.

FEATURES
 source
 1. .3399
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM_DB:9630005B22"
 /db_xref="taxon:10090"
 /clone="9630005B22"
 /tissue_type="cerebellum"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="16 days neonate"
 264..1757
 /note="unnamed protein product; hypothetical Metalloprotease (ADAM type)/reprolysin (M12B) family containing protein (InterPro|IPR001590, evidence: InterPro)
 putative"
 /codon_start=1
 /protein_id="BAC29190.1"
 /db_xref="GI:26330920"
 /translation="MECALLCALRAAGPWPAGLRLAKALQLCCFCASVAV ALASDSSGSGSLNDYVVFVEVDSSGYISHDILHRRKRRAAGASNSLHYRVS AFQDGLHLKLSAILSSHFVQLGKGASRETRPEVPCQLYQGFIRNDSVSVAS TCAGLGLLRTDNFLI SPLQLAQAEHNSYPAGHHPVLYKRTAEKRVWQDYP GSQRTYPGHSPGHTTPASQOPEYSHRWKRFHPCGRKKYAPPPADAYLRDEY GGTGPRRSGKSGNLVETLVADAKMVEKHGKDDTYTLLTVNWSVSLFKDGTI GDSINTVVSLLILLEPEGLINHADSLNSFCQMSALVGNKGRKRDHAILLTGF DICSWNEPCDITGTPAPISGKSCRTINEDTGLGIAFTIAHESGHNFMVHDGEG NPCRKAEGNIMSFITLGNNGVFSWSSCRQYLKFLRTPRCIAFLTG"

CDS

ORIGIN
 Alignment Scores:
 Pred. No.: 4,97e-104 Length: 3399
 Score: 1284.50 Matches: 263
 Percent Similarity: 64.94% Conservative: 63
 Best Local Similarity: 52.39% Mismatches: 119
 Query Match: 48.88% Indels: 58
 DB: 3 Gaps: 10

US-10-804-457-2 (1-491) x AK035797 (1-3399)

Qy 25 AlaProAlaCysAlaMetGlyProAlaAlaAlaProGlySerProSerValProArg 44
 Db 271 GGGCCCTCTCTGT-----GCTTGTGGCCCTCCGGGCGCAGGTCCCGGG- 314
 Qy 45 ProProProProAlaGluArgProGlyTrpMetGluLysGly----- 58
 Db 315 CCGCGGTGGGGCGCGCGGACTGGGGCGTCTGCGCAAGGCACCTCCAGCTGTGCTGCTTC 374
 Qy 58 ----- 58
 Db 375 TCCTGTGCTTCGTGGCGCGTCCCTTACCGAGTCACACGCGCAGCGGTGGCAGCGGA 434
 Qy 59 -----GlutyrAspLeuValSerAlaTyrGluValAlaPheHisArgGlyAspTyrVal 75
 Db 435 TTAATGATGATTACGTCTTTGTGTCGCCAGTAGAGTGGACTCGGCGCGGTCAATATT 494

QY 76 SerHisGluIleMetHisGlnArgArgArgAlaValSerGluValGlu 95
 DB 495 TCACGACGATTTTGCACACAGGAAGGCGACACGGTGCCAGC-----AAC 548
 QY 96 SerLeuHisLeuArgLeuLeuysGlySerArgHisAspPheHisValAspLeuArgThrSer 115
 DB 549 TCCCTGCACACGAGTTTCAGCATTTGACAGGATCTACATTTAGAGCTTAAG---CCC 605
 QY 116 SerSerLeuValAlaProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSer 135
 DB 606 TCGGCGATCTTGACGAGTCACCTTTAGAGTCCAAAGTACTTGGAAAGACGGTGCTTCA--- 662
 QY 136 ValGlnThrLeuProProGlu---AspPheCysPheTyrGlnGlySerLeuArgSerHis 154
 DB 663 ---GAGACTCGGGAGCGCTGAGGTACCGCAGTGTGTATCAGGGATTTATCAGAAATGAC 719
 QY 155 ArgAsnSerSerValAlaLeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGlu 174
 DB 720 AGCTCGCTCTGTGGCTGTGTCTAGCTGTCCGGCTGTGTGAGTCTGTATAGGACCGCA 779
 QY 175 GluAlaAspTyrPheLeuArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAla 194
 DB 780 GACAATGAATTCCTCATCTCACCGCTACCTCAGCTGTGGCCCGACACCAACTATAGC 839
 QY 195 AlaGlnGlySerSerProSerHisValLeuTyrLysArgSerThrGluProHisAlaPro 214
 DB 840 TCGCCTGCGAGGCCACCATCCTCAGCTCTGTGTACAAAGACGACGAGAGGGTTCAGA 899
 QY 215 GlyAlaSerGluValLeuValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHis 234
 DB 900 TGGTACCAAGACTACCTCGCTGCCCGCGGACATATCCT---GGTCACCTCCCAAGTCA 956
 QY 235 Ser-----SerAspLeuArgLeuGlyLeuPro 243
 DB 957 ACTCCCTCCCTCCAGAGCCAGAGCCAGAGTACAGCCATCGAAGTGG----- 1007
 QY 244 GlnLysGlnHisPheCysGlyArgArgLysTyrMetProGlnProProLysGluAsp 263
 DB 1008 CAAAAGCGCATTTTGTGGAGCGACCAAGAATAATATGCCCCCAAGCTCTCGCAGAGGAC 1067
 QY 264 LeuPheIleLeuProAspGluTyrLysSerCysLeuArgHisLeuArgSerLeuLeuArg 283
 DB 1068 GCCTATCTACGCTTTGACGAATATGAGGCGACAGGGCGCCAGAGCGGTCACTGGAAAG 1127
 QY 284 SerHisArgAsnGluLeuLeuAsnValGluThrLeuValValAspLysLysMetMet 303
 DB 1128 TCACAAATAGT-----CTCAATGTGGAGACCTTGTGTGGCAGATCCCAAGATGGTG 1181
 QY 304 GlnAsnHisGlyHisGluAsnIleThrTyrValLeuThrIleLeuAsnMetValSer 323
 DB 1182 GAGAAGCAGCGCAAGGATGACGCTCACCGTACATCTTCACAGTCATGTAACATGGTTTCT 1241
 QY 324 AlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAsnIleValGlyLeuIle 343
 DB 1242 AGCCTGTTCAAGATGGGACCATTTGAAGCGCATATAAATTGTGTGTGAGCCTTAAT 1301
 QY 344 LeuLeuGluAspGluGlnProGlyLeuValIleSerHisHisAlaAspHisThrLeuSer 363
 DB 1302 CTGCTGGAGAGAACCTGAAGGATGCTGTATCAATCACCATGCAGACAGTCTCTAATAC 1361
 QY 364 SerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHisAspHisAla 383
 DB 1362 AGCTTCTGTGAGGAGTTCACATTTGTTGGGAAAGATGGCAAGAGACACGACACGCGC 1421
 QY 384 IleLeuLeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAspThrLeuGly 403
 DB 1422 ATCCCTCTCACAGGATTTGACATTTGTTCTGGAGAACGAAACCATGTGACACACTAGGA 1481
 QY 404 PheAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAsnGluAspThr 423
 DB 1482 TTTGCTCTCTATCAGTGGCATGTGCAGTAGTACCGAGAGCTGTACCATCATGAGACACA 1541
 QY 424 GlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAsnPheGlyMetIleHis 443

DB 1542 GGACTTGGCTTGGCCTTCCACATTGGCATAGTCAGGCGCAACATTTTGGCATGGTTCAT 1601
 QY 444 AspGlyGluGlyAsnMetCysLysLysSerGluGlyAsnIleMetSerProThrLeuAla 463
 DB 1602 GATGTGAAGCAATCCCTCGAGGAAGCAGAGAACAACATCATGTCCACACACTGACT 1661
 QY 464 GlyArgAsnGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHisLysPheLeu 483
 DB 1662 GGAACAATGGGGTGTCTTTTCATGGTCTGCTGCGAGCCGACAGTATCTAAAGAAATTCCTC 1721
 QY 484 ArgSer 485
 DB 1722 AGAAT 1727
 RESULT 4
 LOCUS AK053851
 DEFINITION Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone: E130314N14 product: hypothetical Metalloprotease (ADAM type)/reprolysin (M12B) family containing protein, full insert sequence.
 ACCESSION AK053851
 VERSION AK053851.1 GI:26343798
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 Carninci, P. and Hayashizaki, Y.
 TITLE High efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Iehii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, S., Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 6 (bases 1 to 4335)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,

Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saigo, R., Saichou, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayaehizaki, Y.
 Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-3222, Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.jp/
 URL: http://fantom.gsc.riken.jp/.

FEATURES

source

```

1. 4335
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:E130314N14"
/db_xref="taxon:10090"
/clone="E130314N14"
/tissue_type="eyeball"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
992..4270

```

CDS

```

/notes="unamed protein product; hypothetical
Metalloprotease (ADAM type)/reprolysin (M12B) family
containing protein (InterPro|IPR001590, evidence:
InterPro)
putative"
/codon_start=1
/protein_id="BAC35556.1"
/db_xref="GI:26343799"
/cranlation="MNSSHRYLSCWPRNTTARLQATILTSCTKGQQRSGDGTKT
LAPSGHILVPTQVTLPPRAKQSTAEKGGKGFVDDARNVYAPKPADAYLRPD
EYGTGRPRRSAGKQNLNVLTVADAKVXKHGKDDVTYTLTWNNVSLFKDG
TIGSDINIWVSLILLEPEGLLNHHADQSLNSFCOMQSAALVKGKRRHDAIILT
GFDICSWKNEPCDTLGPISCMKSKYRSCSTINEDTGLAFTIAHSGHNFVGHG
EENPCRKAEGNIMSTLTGNNGVFSWSSCSRYLKKFLUSTPOAGCLVDEPKQTQYKY
PKLPGQYLDAMQCKQWFGAKKLCISGVKMDICKLWCHRVGRHRCETKFPMAEET
ACGLSMWCRQGCQVKGELGPRPIHGWMSAKWSECSRTCGGVKQFERHCSNPKPO
YGGKYCPGSRYSYKLCNPNPCNSLDPRACQCAEYNNKPPRGLYRWKPYTKVEED
RCKLYKAEENFEFFPMGKVKDQTPCSPHRNDVICDGI CELVSCDHELGSKAVSDAC
KYLGKDNSTCKFYKGLYLSQHKANEVYPVVTIPAGARSIETQELQLSSYLAVRSLQ
KYLNGWSIDWPGDFTAGTTFEQRFSFNERLYAPGPNTELVEILLQGNNGPI
AKWALPKMANTQPATRKHTWRTVQSDCVTCGGYISIKALICLRDQHTQVNSGF
CSVRTKATEPKINAFSCPAYLWPGWSACSKSCAGGQSRKIRCKQKPFQKEEAV
LHSLCPVPTQVQVNSHACPPENSPSQCSKTCTGRVRRREVLCQKPAATLEP
SLCSSPRPEACGVLGRCPNNRLQWIASAGSECSATCGLGVKRELKCVKTELQG
KLITPPERCRNIKKAKPGTCRSLQKDLPSVHGGLVFTIWTAVHLSNLRGSPDS
LCPAACSKAGLPQVCSIRSLRCSEPTVPTSVQLKRWTHLWISAGVWNSLSTES
ATTNPMESSAVDHAQGRADLGVSSPSSDDUTRFSUGLIRIKTTHNRGNPIGFKEWTD
HSPS"

```

ORIGIN

Alignment Scores:
 Pred. No.: 4,96e-101 Length: 4335
 Score: 1253.00 Matches: 263
 Percent Similarity: 61.39% Conservative: 63
 Best Local Similarity: 49.53% Mismatches: 122
 Query Match: 47.68% Indels: 83
 Db: 3 Gaps: 12

```

US-10-804-457-2 (1-491) x AK053851 (1-4335)
QY 25 AlaProAlaCysAlaMetGlyProAlaAla---AlaAlaProGlySerProSerValPro 43
Db 376 GGGCCCTCTGTGCTTGTGGCCCTTCCGGGGCGCAGGTCCCGGGCGGCGGCGGCGGCGG 435
QY 44 ArgPro-----ProProAlaGluArgProGlyTrpMet 55
Db 436 CGGGACTGGGGGCTCTGGCCAAAGGAGGTGCCACCCGAGGGGGGGCCGCACAGCGG 495
QY 56 GluLysGly----- 58
Db 496 GGACGAGGAGGCGGCCCGCCCGCGGACGCTGCTCAGCGGGGGTGTCAACCTTTTTC 555
QY 58 ----- 58
Db 556 TTGCAGGCACTCCAGCTGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 615
QY 59 -----GluTyrAspLeuValSerAlaTyr 66
Db 616 GACAGCGCAGCAGCGGTGGCAGCGGATTAATGATGATTACGTCTTTGTCGTCAGTA 675
QY 67 GluValAspHisArgGlyAspTyrValSerHisGluIleMetHisHisGlnArgArg 86
Db 676 GAAGTGGACTCGGGCGGTCTATATTTTACACACATTTTGCACCCACACAGGAAAGCGCA 735
QY 87 ArgAlaValAlaValSerGluValGluSerLeuHisLeuArgLeuLysGlySerArgHis 106
Db 736 TCGGCACACGGTGCAGC-----AACTCCTCGCACTACCGAGTTTTCAGCATTTGGACAG 789
QY 107 AspPheHisValAspLeuArgThrSerSerSerLeuValAlaProGlyPheIleValGln 126
Db 790 GATCTAGCATTTAGAGCTTAAG---CCCTCGCGCATTTTGAGCAGTCACTTTAGAGTCAA 846
QY 127 ThrLeuGlyLysThrGlyThrLysSerValGlnThrLeuProProGlu---AspPheCys 145
Db 847 GTACTTGGAAAGGAGCGTGTCTCA-----GAGACTCGGAGCGCTGAGGTACCGCAGTGT 900
QY 146 PheTyrGlnGlySerLeuArgSerHisArgAsnSerValAlaLeuSerThrCysGln 165
Db 901 TTGTATCAGGATTTATCAGAAATGACAGCTCGTCTCTGTGGTGTGTCTACGTGTGCC 960
QY 166 GlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPheLeuArgProLeuProSer 185
Db 961 GGCTTGTCAAGTCTGATAGACCCGAGACAATGATCTCTCATCTCACCCTACCTCAG 1020
QY 186 HisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSerProSerHisValLeuTyr 205
Db 1021 CTGCTGCCCCAGGAACAACAATACTATAGTCTGCCCTGCAGGCCACCATCTCAGTCTGTAC 1080
QY 206 LysArgSerThrGluProHisAlaProGlyAlaSerGluValLeuValThrSerArgThr 225
Db 1081 AAAAGGACACGACGAGAGGGTCAAGTGGTACCAAGACTACCTCGCTCCACGCGGACA 1140
QY 226 TrpGluLeuAlaHisGlnProLeuHisSer----- 235
Db 1141 TATCCT---GGTCACTCCCCCAAGTCACTCCCTCCCTCCAGAGCCAGAGCCAGAG 1197
QY 236 ---SerAspLeuArgLeuGlyLeuProGlnLysGlnPheCysGlyArgArgLysLys 254
Db 1198 TACAGGCATCGAAGGTGG-----CAAAAGCGGCATTTTGTGGACGACGCAAGAAA 1248
QY 255 ---TyrMetProGlnProProLysGluAspLeuPheIleLeuProAspGluTyrLysSe 273
Db 1249 TGTATATGCCCCCAAGCCTCTCTGACAGAGGACGCTATCTATCGCTTTGACGAATATGGAGG 1308
QY 273 rCysLeuArgHisLysArgSerLeuLeuArgSerHisArgAsnGluGluLeuAsnValGln 293
Db 1309 CACAGGGCGGCCACAGCAGGTGAGTGGAAAGTCACAAAATGGT-----CTCAATGTGA 1362
QY 293 uThrLeuValValValAspLysLysMetMetGlnAsnHisGlyHisGluAsnIleThrTh 313
Db 1363 GACCTTGTGTGGCAGATGCCAAGATGTTGGAGAACGCGGCAAGGATGACGTCAACAC 1422

```



```

QY 313 rTyrValLeuThrIleLeuAsnMetValSerAlaLeuPheLysAspGlyThrIleGlyG1 333
Db 1423 GTACATTCTCACAGTCATGAACATGGTTTCTAGCCTGTTCAAGATGGGACCATTTGAAG 1482
QY 333 yAsnIleAsnIleAlaIleValGlyLeuIleLeuLeuGluAspGluGlnProGlyLeuVa 353
Db 1483 CGACATAAACATTGGTGTGTGAGCCCTAAATCTGTGGAAAGAACCTGAGGATGCT 1542
QY 353 lIleSerHisHisAlaAspHisThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMe 373
Db 1543 GATCAATCACCATGCGACGACGAGCTCTAAACAGCTTCTGTCACTGGCAGTCAGCTCTGT 1602
QY 373 tGlyLysAspGlyThrArgHisAspHisAlaIleLeuLeuThrGlyLeuAspIleCysSe 393
Db 1603 GGGAAAGAATGGCAAGAGACACACGACGACCTCTCTCAGAGGATTTGACATTTGTTTC 1662
QY 393 rTTPlysAsnGluProCysAspThrLeuGlyPheAlaProIleSerGlyMetCysSerLy 413
Db 1663 CTGGAAGAAGCAACCATGTGACACACTAGGATTTGCTCTCATAGTGGCAGTGGCAGTAA 1722
QY 413 sTyrArgSerCysThrIleAsnGluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHi 433
Db 1723 GTACCGAGCTGTACCATCAATGAACACAGGACTTGGCTGGCTTCCACCATTTGGCA 1782
QY 433 sGluSerGlyHisAsnPheGlyMetIleHisAspGlyGluGlyAsnMetCysLysLysSe 453
Db 1783 TGAGTCAGGCGCACACTTTGGCATGTTCTATGATGGTGAAGGCAATCCCTCGCAGGAAGC 1842
QY 453 rGluGlyAsnIleMetSerProThrLeuAlaGlyArgAsnGlyValPheSerTrpSerPr 473
Db 1843 AGAAGGCAACATCATGTGACCCACACTGACTGGAACAAATGGGGTGTTCATGGTGGTC 1902
QY 473 oCysSerArgGlnTyLeuHisLysPheLeu 483
Db 1903 CTGCAGCCGACAGTACTTAAGAAATTCCTC 1933

RESULT 5
CD355731
LOCUS
DEFINITION
UI-M-FYO-cgo-o-23-0-UT.r1 NIH BMAP_FY0 Mus musculus cDNA clone
IMAGE:30355654 5', mRNA sequence.
CD355731
CD355731.1 GI:31148232
EST.
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefi.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5,
Location/Qualifiers
1. 728
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="C57BL/6"
/stable_id="10090"
/clone="IMAGE:30355654"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"

```

FEATURES

source

ORIGIN

```

Alignment Scores:
Pred. No.: 2,74e-86 Length: 728
Score: 1076.00 Matches: 201
Percent Similarity: 89.26% Conservative: 15
Best Local Similarity: 83.06% Mismatches: 26
Query Match: 40.94% Indels: 0
DB: 6 Gaps: 0

US-10-804-457-2 (1-491) x CD355731 (1-728)
QY 186 HisLeuSerTrpIleLeuGlyArgAlaAlaGlnGlySerSerProSerHisValLeuTyr 205
Db 3 CACCTAACAGCAAACTGAACAGATCTGCACAGGGGACTCCCTCCCTCCCATGTTCTGTAC 62
QY 206 LysArgSerThrGluProHisAlaProGlyAlaSerGluValLeuValThrSerArgThr 225
Db 63 AGAGATCTACAGAGCGCCAGGCTCTAGAGAAACGAGTCTCTGATGATCACCAGGAAG 122
QY 226 TrpGluLeuAlaHisGlnProLeuHisSerSerAspLeuArgLeuGlyLeuProGlnLys 245
Db 123 CGAGATCTGGCAAGACACACCTGCACCATGACAACTTCATCTTGGCCCTCACCAGAAAG 182
QY 246 GlnHisPheCysGlyArgGlyValSerTyrMetProGlnProPolysGluAspLeuPhe 265
Db 183 CAGCATTTCTGTGGAAGACGCAAGAAATACATCCCTCAGCTCCCAACAGTACCTCTAC 242
QY 266 IleLeuProAspGluTyrLysSerCysLeuArgHisLysArgSerLeuLeuArgSerHis 285
Db 243 ATCTTGCTGTGAGTACAGGCTAGTTCGCGACACAGGCTCGCTCTTGAATCCAC 302
QY 286 ArgAsnGluGluLeuAsnValGluThrLeuValValAspLysLysMetMetGlnAsn 305
Db 303 AGAAATGAAGAGTTAAATGTGGAGACACTGGTGTGGACAGAAAGATGATGACAAAGC 362
QY 306 HisGlyHisGluAsnIleThrThrValLeuThrIleLeuAsnMetValSerAlaLeu 325
Db 363 CATGGCCATGAAACAACTAGCATTTATGCTCCTCACCATTCTCAACATGGTGTCTGCTTA 422
QY 326 PheLysAspGlyThrIleGlyLysAsnIleAlaIleValGlyLeuIleLeuLeu 345
Db 423 TTCNAGATGGAAACATTTGGGGAAACATCACAATTTGTCATTGTGGGACTCATTCTGCTG 482
QY 346 GluAspGluGlnProGlyLeuValIleSerHisHisAlaAspHisThrLeuSerSerPhe 365
Db 483 GAAGATGAACAGCAGGCTGGCGATTAGTACCATGACGACACACGCTTAAACAGCTTC 542
QY 366 CysGlnTrpGlnSerGlyLeuMetGlyLysAspGlyThrArgHisAspHisAlaIleLeu 385
Db 543 TGCAGTGGCAATCCCGATTTGATGGGAAAGATGGAACTCGTCATGACACACGCACTCTTA 602
QY 386 LeuThrGlyLeuAspIleCysSerTrpLysAsnGluProCysAspThrLeuGlyPheAla 405
Db 603 CTGACTGGTCTGGACATATGTTCTCTGGAAGATGAGCCCTGTGACACATTTGGGTTTGA 662

```

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH_BMAP_FY0"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bontalio, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGACAGC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

Qy 406 ProfileSerGlyMetCysSerLysTyrArgSerCysThrIleAsnGluAspThrGlyLeu 425
 Db 663 CCCAAGCGGATGTCAGTAAGTACCGCAGCTGCACAGTCAATNAGACTCAGGACTT 722
 Qy 426 GlyLeu 427
 Db 723 GGACTG 728
 RESULT 6
 AK036137
 LOCUS
 DEFINITION
 Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length
 enriched library, clone:9630038L21 product:hypothetical
 Metalloprotease (ADAM type)/reprolysin (M12B) family containing
 protein, full insert sequence.
 AK036137
 AK036137.1 GI:26085046
 HTC; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 4593)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hirao, T., Hirozane, T.,
 Horii, F., Imoto, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T.,
 Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ono, M., Oheato, N.,
 Okazaki, Y., Saito, R., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.

TITLE
 JOURNAL
 Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail:genome-res@gscc.riken.jp,
 URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
 Fax:81-45-503-9216)
 CNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL:http://genome.gsc.riken.jp/
 URL:http://fantom.gsc.riken.jp/
 FEATURES
 source
 1. .4593
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strains="C57BL/6J"
 /db_xref="FANTOM_DB:9630038L21"
 /db_xref="taxon:10090"
 /clone="9630038L21"
 /tissue_type="cerebellum"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="16 days neonate"
 misc_feature
 1. .4593
 /note="hypothetical Metalloprotease (ADAM type)/reprolysin
 (M12B) family containing protein (InterPro|IPR001590,
 evidence: InterPro)"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.21e-80 Length: 4593
 Score: 1027.00 Matches: 208
 Percent Similarity: 69.61% Conservatice: 44
 Best Local Similarity: 57.46% Mismatches: 68
 Query Match: 39.08% Indels: 43
 DB: 3 Gaps: 6
 US-10-804-457-2 (1-491) x AK036137 (1-4593)
 Qy 160 AlaleuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPhe 179
 Db 1 GCTGTGCTACGTTGTCGGCTTGTCAAGTCTGTAAGGACCCGACCAATGAATCTCTC 60
 Qy 180 LeuArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSer 199
 Db 61 ATCTCACCCTACCTCAGCTGCTGGCCAGGAACACAACTAGTCTGGCTCGAGCCAC 120
 Qy 200 ProSerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluVal 219
 Db 121 CATCTCAGCTCTGTACAAAGGACAGCAGAGAGGGTCAGATGGTACCAAGACTAC 180
 Qy 220 LeuValThrSerArgThrTrpGluLeuAlaHisGlnProLeuHisSer----- 235
 Db 181 CTGGTCTCCAGCGGACATATCTCT---GGTCACTCCCAAGTCACACTCCCTCCCTCC 237
 Qy 236 -----SerAspLeuArgLeuGlyLeuProGlnLysGlnHisPhe 248
 Db 238 CAGAGCCAGAGCCAGAGTACAGCCATCGAAGGTGG-----CAAAAGCGCATTTT 288
 Qy 249 CysGlyArgArgLysLys-----TyrMetProGlnProProLysGluAspLeuPheIleLe 267
 Db 289 TGTGACGACGACGCAAGAAATGATATATGCCCCCAAGCTCTCTCGAGGAGCGCTATCTACG 348
 Qy 267 uProAspGluTyrLysSerCysLeuArgHisLysArgSerLeuLeuArgSerHisArgAs 287
 Db 349 CTTTGACCAATATGGAGGCACAGGCGGCCGACAGCGGTACGCTGCAAAAGTCACAAATGG 408
 Qy 287 nGluLeuLeuAsnValGluThrLeuValValAspLysLysMetMetGlnAsnHisG 307

```

Db      409 T-----CTCAATGTGGAGACCTTGTGTGGCAGATGCCAAGATGGTGGAGAGCACGG 462
QY      307 yHisGluAenileThrTyrValLeuThrIleLeuAenMetValSerAlaLeuPhely 327
Db      463 CAAGATGACGTCACACGCTACATCTCTCAGCATGAACATGGTTCTAGCCTGTTCAA 522
QY      327 sAepGlyThrIleGlyGlyAenileAenileAlaIleValGlyLeuIleLeuGluAs 347
Db      523 AGATGGGACCATTTGGAAGGACATAAACATTTGTGTGTGGCTTAATCTCTGTGGAGA 582
QY      347 pGlu----- 348
Db      583 AGAAGCTTACAAAGGCTCAACCTTCTCTCTCTGSCACATCATCACTCTTTAATCCCT 642
QY      349 -----GlnProGlyLeuValIleSerHisHisAlaAAspHisThrIleuSerSe 364
Db      643 GCACCTTTCCCATAGGAA--GGACTGCTGATCAATCAATCAATGAGACAGCAGTCTCTAAACAG 701
QY      364 rPheCysGlnTrpGlnSerGlyLeuMetGlyIleAAspGlyThrArgHisAAspHisAlaI 384
Db      702 CTTCTGTGAGTGGCAGTCACTCTTGTGGAAAGATGGCAAGAGACAGCACCCACCCAT 761
QY      384 eLeuLeuThrGlyLeuAAspIleCysSerTrpLysAenGluProCysAAspThrIleuGlyPh 404
Db      762 CTCCTCACAGGATTTGACATTTGTTCTTGGAAAGACCAATGATGACACACTAGGAT 821
QY      404 eAlaProIleSerGlyMetCysSerLysTyrArgSerCysThrIleAenGluAAspThrGl 424
Db      822 TGTCTCTATCAGTGGCATGTGCAGTAAGTACCGAAGCTGTACCATCAATGAAGACACAGG 881
QY      424 yLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHisAAspGlyMetIleHisAs 444
Db      882 ACTTGGCCTTGGCCTTCACAAATTCGATGAGTCAGGAGCAACATTTGGCATGTGTTCA 941
QY      444 pGlyGluGlyAenMetCysLysLysSerGluGlyAenIleMetSerProThrIleuAlaGl 464
Db      942 TGGTGAAGGCAATCCCTCGAGAAAGACGAAGGCAACATCATGTCAACCCACACTGACTGG 1001
QY      464 yArgAenGlyValPheSerTrpSerProCysSerArgGlnTyrLeuHisLysPheLeuAr 484
Db      1002 AAACATGGGTGTTTTCATGTGTCCTGCGAGCCGACAGTATCTAAAGAAATTCCTCAG 1061
QY      484 gSer 485
Db      1062 GACA 1065

RESULT 7
BM845406
LOCUS
DEFINITION
BM845406 537 bp mRNA linear EST 06-MAR-2002
K-EST0123712 S12SNU216 Homo sapiens cDNA clone S12SNU216-99-B04 5',
mRNA sequence.
ACCESSION
BM845406.1 GI:19201805
VERSION
EST.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 537)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
TITLE
JOURNAL
COMMENT
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongseung@mail.kribb.re.kr
Plate: 99 row: B column: 04
High quality sequence stop: 537.

```

FEATURES

Location/Qualifiers
 1..537
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S12SNU216-99-B04"
 /sex="F"
 /tissue_type="Lymph node"
 /cell_type="Epithelial"
 /cell_line="SNU-216"
 /lab_host="Top10F"
 /clone_lib="S12SNU216"
 /note="Organ: Stomach; Vector: pcNS; Site_1: EcoRI;
 Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tobacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F' by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."

ORIGIN

Alignment Scores:
 Pred. No.: 2,35e-72 Length: 537
 Score: 919.00 Matches: 172
 Percent Similarity: 98.31% Conservative: 3
 Best Local Similarity: 96.63% Mismatches: 2
 Query Match: 34.97% Indels: 1
 DB: 4 Gaps: 0

US-10-804-457-2 (1-491) x BM845406 (1-537)

```

QY      315 ValLeuThrIleLeuAenMetValSerAlaLeuPheLysAAspGlyThrIleGlyGlyAen 334
Db      1 GTGCTCACGATACTCAACATGGTATCTGCTTATTCAAAGATGGAACAATAGGAGGAAC 60
QY      335 IleAenIleAlaIleValGlyLeuIleLeuLeuGluAAspGluGlnProGlyLeuValIle 354
Db      61 ATCAACATTTGCAATTTGATGCTCTGATTTCTTAGAAGATGAACAGCAGGACTGGTATA 120
QY      355 SerHisHisAlaAAspHisThrIleuSerSerPheCysGlnTrpGlnSerGlyLeuMetGly 374
Db      121 AGTCACACGACAGACACACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGG 180
QY      375 LysAAspGlyThrArgHisAAspHisAlaIleLeuLeuThrGlyLeuAAspIleCysSerTrp 394
Db      181 AAAGATGGGACTCGTCATGCCACCCATCTTACTGACTGGTCTGGATATATATTTCTCTGG 240
QY      395 LysAenGluProCysAAspThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyr 414
Db      241 AGATATGACCCCTGTGACACTTTGGGATTTGACCCCATTAAGTGAATGTAGTAATAT 300
QY      415 ArgSerCysThrIleAenGluAAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGlu 434
Db      301 CGCAGCTGCACGATTAATGAAGATACAGGCTCTTGACTGGCTTCCCTTACCATTTGCCATGAG 360
QY      435 SerGlyHisAAspPheGlyMetIleHisAAspGlyGluGlyAAsnMetCysLysLysSerGlu 454
Db      361 TCTGGACACAACCTTTGGCATGATTCATGATGAGAAGGGAACATGTGCAAAAGTCCGAG 420
QY      455 GlyAenIleMetSerProThrLeuAlaGlyArgAAsnGlyValPheSerTrpSerProCys 474
Db      421 GGCACATCATGTCTCCCTACATTTGGCAGGACCAATGGAGTCTTCTCTCTGGTCAACCTGC 480
QY      475 SerArgGlnTyrLeuHisLysPheLeuArgSer-VallysMetProAlaLeu 491

```


Site_2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10⁸ by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Alignment Scores:
Pred. No.: 2,97e-72 Length: 625
Score: 919.00 Matches: 172
Percent Similarity: 98.31% Conservative: 3
Best Local Similarity: 96.63% Mismatches: 2
Query Match: 34.97% Indels: 1
DB: 4 Gaps: 0

US-10-804-457-2 (1-491) x BM850160 (1-625)

QY 315 ValLeuThrIleLeuMetValSerAlaLeuPheLysAspGlyThrIleGlyValAsn 334
DB 1 GTGCTCAGATATCAACATGGTATCTGCTTATTCAGAGTGGACATAGAGGAAAC 60
QY 335 IleAsnIleAlaIleValGlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIle 354
DB 61 ATCAACATTCGAATTTAGTCTGATTTCTTCTAGAGATGAACAGCAGGACTGGTATA 120
QY 355 SerHisHisAlaSerHisThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGly 374
DB 121 AGTCACACGACGACACACCTTAAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGG 180
QY 375 LysAspGlyThrArgHisAspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrp 394
DB 181 AAAGATGGAGCTCTGTCATGCCAGCGCATCTTACTACTGCTGGTCTGATATATGTTCTGG 240
QY 395 LysAsnGluProCysAspThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTrp 414
DB 241 AAGAATGAGCCCTGTGACACTTTGGGATTTGCACCCATAAGTGAATGTGTAGTAAATAT 300
QY 415 ArgSerCysThrIleAsnGluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGlu 434
DB 301 CGCAGCTGCACGATTAATGAAGATACAGGCTTTGGACTGGCCCTTCACCATTTGCCATGAG 360
QY 435 SerGlyHisAsnPheGlyMetIleHisAspGlyGluGlyAsnMetCysLysLysSerGlu 454
DB 361 TCTGGACACACTTTGGCATGATTCATGATGGAGAGGGACATGTGCAAAAGATCCGAG 420
QY 455 GlyAsnIleMetSerProThrLeuAlaGlyArgAsnGlyValPheSerTrpSerProCys 474
DB 421 GGCAACATCATGTCCCTCATTTGGCAGGACGCAATGGAGTCTTCTCTGGTCAACCTGC 480
QY 475 SerArgGlnTrpLeuHisLysPheLeuArgSerValLysMetProAlaLeu 491
DB 481 AGCGCCGAGTATACACAAATTTCTAAAGCACCCTCAAGCTATCTGCCCTTG 532

RESULT 10

CD636687/c
LOCUS 56008668H1 FLP Homo sapiens cDNA, mRNA linear EST 12-JAN-2004
DEFINITION CD636687
ACCESSION CD636687.1 GI:40284954
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

1 (bases 1 to 523)
AUTHORS Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
TITLE Circular rapid amplification of cDNA ends for high-throughput extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
COMMENT Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com

FEATURES

source

Location/Qualifiers
1..523
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN

Alignment Scores:
Pred. No.: 5,96e-69 Length: 523
Score: 881.00 Matches: 168
Percent Similarity: 98.83% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 2
Query Match: 33.52% Indels: 0
DB: 6 Gaps: 0

US-10-804-457-2 (1-491) x CD636687 (1-523)

QY 60 TyrAspLeuValSerIleAlaTrpGluValAspHisArgGlyAspTrpValSerHisGluIle 79
DB 515 TATGACTCTGCTCTCTCCTACGAGGTGACACAGGGCGGATACGTCCTCCATGAATC 456
QY 80 MetHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeu 99
DB 455 ATGCACCATCAGCGCGGAGAGAGCAGTGCCTGTCCAGAGTTGAGTCTCTTCACTT 396
QY 100 ArgLeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerLeuVal 119
DB 395 CGGCTGAAGGCCCGCAGGCACGACTTCCACATGATCTGAGGACTTCCAGCAGCTAGTG 336
QY 120 AlaProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeu 139
DB 335 GCTCTCTGGCTTATTGTGACAGCGTTGGGAAAGACAGGCACCTAAGTCTGTGCAGACTTTA 276
QY 140 ProProGluAspPheCysPheTrpGlnGlySerLeuArgSerHisArgAsnSerSerVal 159
DB 275 CGCCACAGAGACTTCTGTTTCTATCAAGGCTCTTTGCGATCACACAGAAACTCTCTAGTG 216
QY 160 AlaLeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTrpPhe 179
DB 215 GCGCTTTCAACCTGCCAAGGCTTGTTCAGCATGATACGACAGAGAGGAGGAGGAGTCTTC 156
QY 180 LeuArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaAlaGlnGlySerSer 199
DB 155 CTAAGGCCACTTCTCTTACACCTCTCTGCGAAACTCGGCAGAGCTGCCCAAGCAGCTCG 96
QY 200 ProSerHisValLeuTrpLysArgSerThrGluProHisAlaProGlyAlaSerGluVal 219
DB 95 CCATCCACGATCTGTACAGAGATCCACAGAGCCCCATGCTCTCTGGGCCAGTGAAGTTC 36
QY 220 LeuValThrSerArgThrTrpGluLeuAlaHis 230
DB 35 CTGGTGACCTCAAGGACATGGGAGCTGGCAGCAT 3

RESULT 11

CD636689/c
LOCUS 56008676H1 FLP Homo sapiens cDNA, mRNA linear EST 12-JAN-2004
DEFINITION CD636689
ACCESSION CD636689.1 GI:40284956
VERSION EST.
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 523)
TITLE Circular rapid amplification of cDNA ends for high-throughput
JOURNAL extension cloning of partial genes
COMMENT Genomics 84 (1), 205-210 (2004)
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
Location/Qualifiers
1. .523
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN
Alignment Scores:
Pred. No.: 5,968-69 Length: 523
Score: 881.00 Matches: 168
Percent Similarity: 98.83% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 2
Query Match: 33.52% Indels: 0
DB: 6 Gaps: 0
US-10-804-457-2 (1-491) x CD636689 (1-523)
QY 60 TyrAspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIle 79
Db 515 TATGACCTGGTCTCTGCTACGAGTTTCACACAGGGCGGATTACGTGCCATGAATC 456
QY 80 MethHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeu 99
Db 455 ATGCACCATCAGCGCGGAGNAGCAGTCCCATGATCTCGAGGTTGAGTCTCTTCACCTT 396
QY 100 ArgLeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerLeuVal 119
Db 395 CGGCTGAAGGCCCGCCAGGACGACTCCCATGATCTCGAGGACTTCAGAGCCTAGTG 336
QY 120 AlaProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeu 139
Db 335 GCTCTGCTTATTTGTGCAGAGCTTGGGAAAGACAGGCACCTAAGTCTGTGCAGACTTTA 276
QY 140 ProProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerVal 159
Db 275 CCGCCAGAGACTTCTGTTCTATCAAGGCTCTTTGCGATCACACAGAACTCTCTCAGTG 216
QY 160 AlaLeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPhe 179
Db 215 GCCTTTCAACCTGCCAGGCTTGTGAGCATGATACGAAAGAGAGCGCGGATTACTTC 156
QY 180 LeuArgProLeuProSerHisLeuSerTrpLysLeuGlyValArgAlaGlnGlySer 199
Db 155 CTAAGGCCACTTCTTCACACCTCTCATGGAACCTCGCAGAGCTGCCAAGGCGAGCTCG 96
QY 200 ProSerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluVal 219
Db 95 CCATCCACGTACTGTACAGAGATCCACAGAGCCCCCATGCTCTCTGGGGCCAGTGAGTTC 36
QY 220 LeuValThrSerArgThrTrpGluLeuAlaHis 230
Db 35 CTGGTGACCTCAAGGACATGGGAGCTGGCACAT 3

RESULT 12
CD636691/c 523 bp mRNA linear EST 12-JAN-2004
LOCUS 56008692H1 FLP Homo sapiens cDNA, mRNA sequence.

ACCESSION CD636691
VERSION 1 GI:40284958
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 523)
TITLE Circular rapid amplification of cDNA ends for high-throughput
JOURNAL extension cloning of partial genes
COMMENT Genomics 84 (1), 205-210 (2004)
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
Location/Qualifiers
1. .523
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN
Alignment Scores:
Pred. No.: 5,968-69 Length: 523
Score: 881.00 Matches: 168
Percent Similarity: 98.83% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 2
Query Match: 33.52% Indels: 0
DB: 6 Gaps: 0
US-10-804-457-2 (1-491) x CD636691 (1-523)

QY 60 TyrAspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIle 79
Db 515 TATGACCTGGTCTCTGCTACGAGTTTCACACAGGGCGGATTACGTGCCATGAATC 456
QY 80 MethHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeu 99
Db 455 ATGCACCATCAGCGCGGAGNAGCAGTCCCATGATCTCGAGGTTGAGTCTCTTCACCTT 396
QY 100 ArgLeuLysGlySerArgHisAspPheHisValAspLeuArgThrSerSerLeuVal 119
Db 395 CGGCTGAAGGCCCGCCAGGACGACTCCCATGATCTCGAGGACTTCAGAGCCTAGTG 336
QY 120 AlaProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeu 139
Db 335 GCTCTGCTTATTTGTGCAGAGCTTGGGAAAGACAGGCACCTAAGTCTGTGCAGACTTTA 276
QY 140 ProProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerSerVal 159
Db 275 CCGCCAGAGACTTCTGTTCTATCAAGGCTCTTTGCGATCACACAGAACTCTCTCAGTG 216
QY 160 AlaLeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluGluAlaAspTyrPhe 179
Db 215 GCCTTTCAACCTGCCAGGCTTGTGAGCATGATACGAAAGAGAGCGCGGATTACTTC 156
QY 180 LeuArgProLeuProSerHisLeuSerTrpLysLeuGlyValArgAlaGlnGlySer 199
Db 155 CTAAGGCCACTTCTTCACACCTCTCATGGAACCTCGCAGAGCTGCCAAGGCGAGCTCG 96
QY 200 ProSerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluVal 219
Db 95 CCATCCACGTACTGTACAGAGATCCACAGAGCCCCCATGCTCTCTGGGGCCAGTGAGTTC 36
QY 220 LeuValThrSerArgThrTrpGluLeuAlaHis 230
Db 35 CTGGTGACCTCAAGGACATGGGAGCTGGCACAT 3

RESULT 13

CD636693/c
LOCUS 56008776H1 FLP Homo sapiens cDNA, mRNA linear EST 12-JAN-2004
DEFINITION CD636693
ACCESSION CD636693
VERSION CD636693.1 GI:40284960
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 522)
AUTHORS Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
TITLE Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
COMMENT Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
FEATURES
source
1..522
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/notes="Vector: pDrive Cloning Vector"
ORIGIN
Alignment Scores:
Pred. No.: 1.08e-67 Length: 522
Score: 867.00 Matches: 166
Percent Similarity: 98.25% Conservative: 2
Best Local Similarity: 97.08% Mismatches: 3
Query Match: 32.99% Indels: 0
Gaps: 0
DB: 6
US-10-804-457-2 (1-491) x CD636693 (1-522)
QY 60 TyrAspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIle 79
DB 514 TATGACCCGGTCTCTGCTACGAGGTGACCAAGGGCGGATACGTGTCCTCAATGAATC 455
QY 80 MetHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeu 99
DB 454 ATGCACCATCAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 395
QY 100 ArgLeuGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuVal 119
DB 394 CGGCTGAAAGGCCCCAGGACGACTTCCACATGGATCTGAGGACTTCCAGCAGCTAGTG 335
QY 120 AlaProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeu 139
DB 334 GCTCTGGCTTATTGTGACAGAGCTTGGGAAAGACAGAGAGAGAGAGAGAGAGAGAG 275
QY 140 ProProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerVal 159
DB 274 CCGCCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGATCATCACAGAACTCTCTCAGTG 215
QY 160 AlaLeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluAlaAspTyrPhe 179
DB 214 GCCCTTTCAACTGCGCAAGGCTTGTGAGGATGATACGAAACAGAGAGGAGGAGTATTCT 155
QY 180 LeuArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaGlnGlySerSer 199
DB 154 CTAAAGGCCACTTCTCTTACACACCTCTCATGGAAGACTCGGAGAGCTGCCAAGGACGCTCG 95
QY 200 ProSerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluVal 219
DB 94 CCATCCACGCTACTGTACAGAGATCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 35
QY 220 LeuValThrSerArgThrTrpGluLeuAlaHis 230
|||||

34 CTGGTGACCTCAAGGACATGGGAGCTGGCAAT 2
Db
RESULT 14
CD636688
LOCUS 56008668J1 FLP Homo sapiens cDNA, mRNA linear EST 12-JAN-2004
DEFINITION CD636688
ACCESSION CD636688
VERSION CD636688.1 GI:40284955
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 523)
AUTHORS Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
TITLE Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
COMMENT Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
FEATURES
source
1..523
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/notes="Vector: pDrive Cloning Vector"
ORIGIN
Alignment Scores:
Pred. No.: 1.06e-66 Length: 523
Score: 856.00 Matches: 168
Percent Similarity: 98.26% Conservative: 1
Best Local Similarity: 97.67% Mismatches: 3
Query Match: 32.57% Indels: 1
Gaps: 0
DB: 6
US-10-804-457-2 (1-491) x CD636688 (1-523)
QY 60 TyrAspLeuValSerAlaTyrGluValAspHisArgGlyAspTyrValSerHisGluIle 79
DB 9 TATGACCTGGTCTCTGCTACGAGGTGACCAAGGGCGGATACGTGTCCTCAATGAATC 68
QY 80 MetHisGlnArgArgArgAlaValAlaValSerGluValGluSerLeuHisLeu 99
DB 69 ATGCACCATCAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 128
QY 100 ArgLeuGlySerArgHisAspPheHisValAspLeuArgThrSerSerSerLeuVal 119
DB 129 CGGCTGAAAGGCCCCAGGACGACTTCCACATGGATCTGAGGACTTCCAGCAGCTAGTG 188
QY 120 AlaProGlyPheIleValGlnThrLeuGlyLysThrGlyThrLysSerValGlnThrLeu 139
DB 189 GCTCTGGCTTATTGTGACAGAGCTTGGGAAAGACAGAGAGAGAGAGAGAGAGAGAG 248
QY 140 ProProGluAspPheCysPheTyrGlnGlySerLeuArgSerHisArgAsnSerVal 159
DB 249 CCGCCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGATCATCACAGAACTCTCTCAGTG 308
QY 160 AlaLeuSerThrCysGlnGlyLeuSerGlyMetIleArgThrGluAlaAspTyrPhe 179
DB 309 GCCCTTTCAACTGCGCAAGGCTTGTGAGGATGATACGAAACAGAGAGGAGGAGTATTCT 368
QY 180 LeuArgProLeuProSerHisLeuSerTrpLysLeuGlyArgAlaGlnGlySerSer 199
DB 369 CTAAAGGCCACTTCTCTTACACACCTCTCGTGGAAACTCGGAGAGCTGCCCAAGGACGCTCG 428
QY 200 ProSerHisValLeuTyrLysArgSerThrGluProHisAlaProGlyAlaSerGluVal 219
DB 429 CATCCACGCTACTGTACAGAGATCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 487
|||||


```
Db 456 ----- 456
QY 238 LeuArgLeuGlyLeuProGlnLysGlnHisPheCysGlyArgArgLysLysTyrMetPro 257
Db 457 -----CCC 459
QY 258 GlnProProlLysGluAspLeuPheLeuLeuProAspGluTyrLysSerCysLeuArgHis 277
Db 460 AAGCCTCTCGCAGAGACCCCTATCTACGCTTTGACGAATATGAGGACACAGGGCGGCC 519
QY 278 LysArgSerLeuLeuArgSerHisArgAenGluGluLeuAenValGluThrLeuValVal 297
Db 520 AGACGGTCAGCTGGAAGATCACAAATGGT-----CTCAATGGAGACCCTTGTGTG 573
QY 298 ValAspLysLysMetMetGlnAsnHisGlyHisGluAsnIleThrTyrValLeuThr 317
Db 574 GCAGATGCCAAGATGTGTGAGAGACGCGCAAGGATGACGTCAACACGTACATTCTCACA 633
QY 318 IleLeuAsnMetValSerAlaLeuPheLysAspGlyThrIleGlyGlyAsnIleAsnIle 337
Db 634 GTCATGAACATGTTTCTAGCCTGTTCAAAGATGGGACCATTTGGAAGCGACATAAACATT 693
QY 338 AlaIleValGlyLeuIleLeuLeuGluAspGluGlnProGlyLeuValIleSerHisHis 357
Db 694 GTGGTTGTGAGCCTAATTCTGTGGAAGAAGAACCTGAAGGAGCTGTGTATCAATCACCAT 753
QY 358 AlaAspHisThrLeuSerSerPheCysGlnTrpGlnSerGlyLeuMetGlyLysAspGly 377
Db 754 GCAGACCACTCTTAACACAGCTTCTGTAGTGGCAGTCAGCTCTTGTGGGAAGATGGC 813
QY 378 ThrArgHisAspHisAlaIleLeuLeuThrGlyLeuAspIleCysSerTrpLysAenGlu 397
Db 814 AAGAGACACGACACGCCCATCTCTCTCACAGATTGACATTTGTTCTGGAAGAACGAA 873
QY 398 ProCysAspThrLeuGlyPheAlaProIleSerGlyMetCysSerLysTyrArgSerCys 417
Db 874 CCATGTGACACACTAGGATTTGCTCTATCAGTGGCATGTGCAGTAAGTACCGAAGCTGT 933
QY 418 ThrIleAsnGluAspThrGlyLeuGlyLeuAlaPheThrIleAlaHisGluSerGlyHis 437
Db 934 ACCATCATGAGACACAGGACTTGGCCTGGCTTCACCATTTGCCGATGAGTCAGGGCAC 993
```

Search completed: June 14, 2005, 14:17:22
Job time : 3716 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2005, 16:34:24 ; Search time 604 Seconds.
(without alignments)
9955.823 Million cell updates/sec

Title: US-10-804-457-3
Perfect score: 3675
Sequence: 1 atgaagcccgccgcccggg.....gctctaagtcacactgtga 3675

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/5A COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfileseq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3675	100.0	3675	3 US-09-930-872-3	Sequence 3, Appli
2	3675	100.0	3675	4 US-10-217-774-3	Sequence 3, Appli
3	3675	100.0	4042	3 US-09-930-872-5	Sequence 5, Appli
4	3675	100.0	4042	4 US-10-217-774-5	Sequence 5, Appli
5	1449.8	39.5	1476	3 US-09-930-872-1	Sequence 1, Appli
6	1449.8	39.5	1476	4 US-10-217-774-1	Sequence 1, Appli
7	306.2	8.3	5357	3 US-09-392-184-5	Sequence 5, Appli
8	300.2	8.2	3571	4 US-09-799-451-411	Sequence 411, App
9	295.4	8.0	3377	4 US-09-981-953A-3	Sequence 3, Appli
10	292.2	8.0	3766	4 US-09-981-953A-1	Sequence 3, Appli
11	266.2	7.2	2274	4 US-09-963-791-23	Sequence 23, Appl
12	266.2	7.2	2727	4 US-09-963-791-1	Sequence 1, Appli
13	255	6.9	3218	3 US-09-369-364A-6	Sequence 6, Appli
14	248.8	6.8	3885	3 US-09-369-364A-16	Sequence 16, Appl
15	160.4	4.4	5804	3 US-09-369-364A-12	Sequence 12, Appl
16	157.2	4.3	1317	4 US-09-963-791-21	Sequence 21, Appl
17	157.2	4.3	1770	4 US-09-963-791-11	Sequence 11, Appl
18	144	3.9	2848	3 US-09-369-364A-4	Sequence 4, Appli
19	141.8	3.9	1071	4 US-09-963-791-19	Sequence 19, Appl
20	141.8	3.9	1524	4 US-09-963-791-9	Sequence 9, Appli
21	141	3.8	2625	3 US-09-369-364A-14	Sequence 14, Appl
22	130.8	3.6	954	4 US-09-963-791-15	Sequence 15, Appl
23	130.8	3.6	1407	4 US-09-963-791-15	Sequence 5, Appli
24	128.6	3.5	3636	4 US-09-949-016-5530	Sequence 5530, Ap
25	127.8	3.5	6692	3 US-09-491-522-1	Sequence 1, Appli
26	127	3.5	2450	3 US-09-491-522-2	Sequence 2, Appli
27	123.2	3.4	2450	3 US-09-491-522-9	Sequence 9, Appli

28	123.2	3.4	4580	3 US-09-491-522-8	Sequence 8, Appli
29	123	3.3	3160	4 US-09-963-791-25	Sequence 25, Appl
30	117.6	3.2	2853	4 US-10-009-332-2	Sequence 2, Appli
31	114.2	3.1	3002	3 US-09-369-364A-1	Sequence 1, Appli
32	109.4	3.0	3706	3 US-09-484-970B-58	Sequence 58, Appl
33	109.4	3.0	4676	3 US-09-130-491-1	Sequence 1, Appli
34	108.4	2.9	2184	4 US-09-445-023A-13	Sequence 13, Appl
35	107.8	2.9	3889	4 US-09-568-559-1	Sequence 1, Appli
36	106.2	2.9	2184	4 US-09-445-023A-2	Sequence 2, Appli
37	103.4	2.8	4858	3 US-09-392-184-1	Sequence 1, Appli
38	101.6	2.8	6659	4 US-09-321-987B-1	Sequence 1, Appli
39	99.2	2.7	1642	3 US-09-369-364A-18	Sequence 18, Appl
40	95.4	2.6	3250	3 US-09-122-126B-14	Sequence 14, Appl
41	95.4	2.6	3250	4 US-09-634-286A-14	Sequence 14, Appl
42	95.4	2.6	3250	4 US-10-247-685-14	Sequence 14, Appl
43	93.8	2.6	3638	3 US-09-369-364A-8	Sequence 8, Appli
44	79.6	2.2	4192	3 US-09-122-126B-1	Sequence 1, Appli
45	79.6	2.2	4192	4 US-09-634-286A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-930-872-3
; Sequence 3, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Erididdle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. 6448388el Human Proteases and Polynucleotides Encoding the S
; CURRENT FILING DATE: 2001-08-14
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3675
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-930-872-3

Query Match	100.0%;	Score 3675;	DB 3;	Length 3675;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3675;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGAAGCCCCCGCGCGCGATGGCGGGCTTGGCGGCGCTGTGGATGCTGTGGCGCAG	60	
Db	1	ATGAAGCCCCCGCGCGCGATGGCGGGCTTGGCGGCGCTGTGGATGCTGTGGCGCAG	60	
QY	61	GTGGCGGAGCAGCACCCTGCGTCCATGGGACCGCGAGCGCGCTGGAGGCCCG	120	
Db	61	GTGGCGGAGCAGCACCCTGCGTCCATGGGACCGCGAGCGCGCTGGAGGCCCG	120	
QY	121	AGCGTCCCGGCTCTCTCCACCGCGAGCGCGCGCTGGATGGAAGGCGCAATAT	180	
Db	121	AGCGTCCCGGCTCTCTCCACCGCGAGCGCGCGCTGGATGGAAGGCGCAATAT	180	
QY	181	GACCTGGTCTCTGCTTACAGGTTGACCAAGGGGCGATTCGTGTCCATGAATCATG	240	
Db	181	GACCTGGTCTCTGCTTACAGGTTGACCAAGGGGCGATTCGTGTCCATGAATCATG	240	
QY	241	CACCATCAGCGGGGAGAGACGCTGGCGGTTCGAGGTTGAGTCTTTCACCTTCG	300	
Db	241	CACCATCAGCGGGGAGAGACGCTGGCGGTTCGAGGTTGAGTCTTTCACCTTCG	300	
QY	301	CTGAAGGCTCCAGGCACGACTTCCACGTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT	360	
Db	301	CTGAAGGCTCCAGGCACGACTTCCACGTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT	360	
QY	361	CCTGGCTTTATTGTGTCAGACGTTGGGAAAGACAGGCACTAAAGTCTGTGCAGACTTTACCG	420	

[illegible]

```
QY 2581 GGGACCGAAGACAGCCCTGCGCCAGCCAGCTACACTTGGGCGCAATCGTGGCGCTCTGAG 2640
Db 2581 GGGACCGAAGACAGCCCTGCGCCAGCCAGCTACACTTGGGCGCAATCGTGGCGCTCTGAG 2640
QY 2641 TGTCTCGGTGCTTGGGAGGGGACAGATGACCGTGTGAGAGAGGGCTGTACAGAGACCTG 2700
Db 2641 TGTCTCGGTGCTTGGGAGGGGACAGATGACCGTGTGAGAGAGGGCTGTACAGAGACCTG 2700
QY 2701 AAGTTTCAAGTAATAATATGCTTCTGCAATCCCAAGACAGACCTGTGCAGGGGCTGGTG 2760
Db 2701 AAGTTTCAAGTAATAATATGCTTCTGCAATCCCAAGACAGACCTGTGCAGGGGCTGGTG 2760
QY 2761 CCTTGCNAAGTATCTGGCTGTCTCCAGAGTGTGCTGGGGAACTGGAGTGTCTGCAGT 2820
Db 2761 CCTTGCNAAGTATCTGGCTGTCTCCAGAGTGTGCTGGGGAACTGGAGTGTCTGCAGT 2820
QY 2821 CGGACGTGTGGGGGGTGCACAGAGCGCCCGTGCAGTGCACACGGGGGTGCACTAT 2880
Db 2821 CGGACGTGTGGGGGGTGCACAGAGCGCCCGTGCAGTGCACACGGGGGTGCACTAT 2880
QY 2881 GACTCGGAGCAGTCCCGGCGAGCCTGTGTGCCCTCAGCCTGTCTCCCTCCAGAGGAGGCC 2940
Db 2881 GACTCGGAGCAGTCCCGGCGAGCCTGTGTGCCCTCAGCCTGTCTCCCTCCAGAGGAGGCC 2940
QY 2941 TGCACCTCTCAGAGTGTCCCACTGTGATGAGCGCCGCGGCTGTGGGAGAGTGTCTCAC 3000
Db 2941 TGCACCTCTCAGAGTGTCCCACTGTGATGAGCGCCGCGGCTGTGGGAGAGTGTCTCAC 3000
QY 3001 ACCTGTGGGAGGGGTGGAGGAGCGGCGAGTGGCTGTAAAGAGCAACCCCTCGGCC 3060
Db 3001 ACCTGTGGGAGGGGTGGAGGAGCGGCGAGTGGCTGTAAAGAGCAACCCCTCGGCC 3060
QY 3061 AGAGCGCAGCTGTGCGCCAGCGCTGTCTGCACCTCCGAGCCCAAGCCAGGATGCAATGAA 3120
Db 3061 AGAGCGCAGCTGTGCGCCAGCGCTGTCTGCACCTCCGAGCCCAAGCCAGGATGCAATGAA 3120
QY 3121 GCCTGTCTGTTCAGCGCTGCCCAAGCCCAAGAGCTGCAGTGTGTGTCGCGCTGG 3180
Db 3121 GCCTGTCTGTTCAGCGCTGCCCAAGCCCAAGAGCTGCAGTGTGTGTCGCGCTGG 3180
QY 3181 TCCCACTGTCTGTGATGTGAAAGAGGACACAGAAAGATCTTAAATGTGCTGAA 3240
Db 3181 TCCCACTGTCTGTGATGTGAAAGAGGACACAGAAAGATCTTAAATGTGCTGAA 3240
QY 3241 AAGTATGTTTCTGGAAGTATCGAGAGCTGGCTCAAGAGAGTGTCAATTTGCCGAAG 3300
Db 3241 AAGTATGTTTCTGGAAGTATCGAGAGCTGGCTCAAGAGAGTGTCAATTTGCCGAAG 3300
QY 3301 CCAGAGCTGGAGCTGGAACGTGCTGCGCCCGCTTCCATGCCCCAGGACCCCCCATTT 3360
Db 3301 CCAGAGCTGGAGCTGGAACGTGCTGCGCCCGCTTCCATGCCCCAGGACCCCCCATTT 3360
QY 3361 GCTGTGCGGGACCTCGAGGGGACAGTGGTTGGCTCACCTGTGTCTCAGTGACGGCC 3420
Db 3361 GCTGTGCGGGACCTCGAGGGGACAGTGGTTGGCTCACCTGTGTCTCAGTGACGGCC 3420
QY 3421 AGCTGTGGGGAGGCGTTCAGACAGAGTCCGTGACGTGCTGTGCGGGGCGCGCGGCC 3480
Db 3421 AGCTGTGGGGAGGCGTTCAGACAGAGTCCGTGACGTGCTGTGCGGGGCGCGCGGCC 3480
QY 3481 TCAGGCTGTCTCTGCAACAGAGCTTGGGCTTCCCTGCGCTGTGCAACACTCACTTCTGC 3540
Db 3481 TCAGGCTGTCTCTGCAACAGAGCTTGGGCTTCCCTGCGCTGTGCAACACTCACTTCTGC 3540
QY 3541 CCCATTGCAAGAGAAAGATGCTTCTGCAAGACTACTTCACTGGTGTCTACCTGGTA 3600
Db 3541 CCCATTGCAAGAGAAAGATGCTTCTGCAAGACTACTTCACTGGTGTCTACCTGGTA 3600
QY 3601 CCCACAGCGGATGTGACCAAGTCTACGCAAGCTGTGCAAGCTGTGCAAGCTGTCTCT 3660
Db 3601 CCCACAGCGGATGTGACCAAGTCTACGCAAGCTGTGCAAGCTGTGCAAGCTGTCTCT 3660
```

```
QY 3661 AAGTCCAACCTTGTGA 3675
Db 3661 AAGTCCAACCTTGTGA 3675

RESULT 2
US-10-217-774-3
; Sequence 3, Application US/10217774
; Patent No. 6734007
; GENERAL INFORMATION:
; APPLICANT: Eriddle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. 6734007el Human Proteases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/10/217,774
; PRIOR FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US/09/930,872
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3675
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-217-774-3
```

Query Match 100.0%; Score 3675; DB 4; Length 3675;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3675; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAAGCCCGCGCGCGGATGGGGGCTTGGGCGCGCTGTGGATGCTGCTGGCGCAG 60
Db 1 ATGAAGCCCGCGCGCGGATGGGGGCTTGGGCGCGCTGTGGATGCTGCTGGCGCAG 60
QY 61 GTGGCGGAGCAGGACCTGCGTCGCGCATGGGACCCGCGAGCGCGCGCTGGAGGCCG 120
Db 61 GTGGCGGAGCAGGACCTGCGTCGCGCATGGGACCCGCGAGCGCGCGCTGGAGGCCG 120
QY 121 AGCGTCCCGGCTCTCTCCACCGCGGAGCGCGCGGCTGTGATGAAAGGCGCAATAT 180
Db 121 AGCGTCCCGGCTCTCTCCACCGCGGAGCGCGCGGCTGTGATGAAAGGCGCAATAT 180
QY 181 GACCTGTCTCTGCTTACGAGGTGACACACGGGCGATTACGTGTCCCATGAAATCATG 240
Db 181 GACCTGTCTCTGCTTACGAGGTGACACACGGGCGATTACGTGTCCCATGAAATCATG 240
QY 241 CACCATCAGCGCGGAGAGAGCAGTGGCGGTGTCCGAGTTGAGTCTCTTCACTTCGG 300
Db 241 CACCATCAGCGCGGAGAGAGCAGTGGCGGTGTCCGAGTTGAGTCTCTTCACTTCGG 300
QY 301 CTGAAAGGCTCCAGGCAACAATTCAAGTGTGATCTGAGGACTTCAGGACGCTAGTGGCT 360
Db 301 CTGAAAGGCTCCAGGCAACAATTCAAGTGTGATCTGAGGACTTCAGGACGCTAGTGGCT 360
QY 361 CTGTGGCTTTATGTGCGACGTTGGGAAAGACAGGACCTAAAGTGTGTGAGACTTTACCG 420
Db 361 CTGTGGCTTTATGTGCGACGTTGGGAAAGACAGGACCTAAAGTGTGTGAGACTTTACCG 420
QY 421 CCAGGAGACTTCTGTTTCTATCAAGGCTTTTGGCATCACAGAAACTCTCTCAGTGGCC 480
Db 421 CCAGGAGACTTCTGTTTCTATCAAGGCTTTTGGCATCACAGAAACTCTCTCAGTGGCC 480
QY 481 CTTTCAACTGCGCAAGGCTTGTGAGGATGATACGAAACAGAGAGGCGAGATTACTCTTA 540
Db 481 CTTTCAACTGCGCAAGGCTTGTGAGGATGATACGAAACAGAGAGGCGAGATTACTCTTA 540
QY 541 AGGCGACTTCTCTCACTCTCATGGAACCTGCGGAGGCTGCGGAGGCTGCGGAGCTCGCCA 600
Db 541 AGGCGACTTCTCTCTCACTCTCATGGAACCTGCGGAGGCTGCGGAGGCTGCGGAGCTCGCCA 600
```

QY	601	TCCACGTA	CTGTACA	GAGATCCA	GAGACCC	CATGCTCT	TGGGGCC	AGTGA	GGTCCTG	660
DB										
QY	601	TCCACGTA	CTGTACA	GAGATCCA	GAGACCC	CATGCTCT	TGGGGCC	AGTGA	GGTCCTG	660
DB										
QY	661	GTGACCT	CAAGCA	TGGAGCT	GGCA	CATCA	ACCCCTG	CA	CAGCAG	720
DB										
QY	661	GTGACCT	CAAGCA	TGGAGCT	GGCA	CATCA	ACCCCTG	CA	CAGCAG	720
DB										
QY	721	GGACTGCC	CAAAA	AGCAGC	ATTCT	GTGTGGA	AGCGCA	AGAAATACA	TGCGCC	780
DB										
QY	721	GGACTGCC	CAAAA	AGCAGC	ATTCT	GTGTGGA	AGCGCA	AGAAATACA	TGCGCC	780
DB										
QY	781	AAGAA	GACCTT	TCATCT	TGCGCA	GTAGTATA	AGTCTT	GTGTTA	CGGCAT	840
DB										
QY	781	AAGAA	GACCTT	TCATCT	TGCGCA	GTAGTATA	AGTCTT	GTGTTA	CGGCAT	840
DB										
QY	841	CTTCGAG	GCTCCC	ATAGAAA	TGAAGT	GAAGTCT	TGCGCA	GTAGTATA	AGTCTT	900
DB										
QY	841	CTTCGAG	GCTCCC	ATAGAAA	TGAAGT	GAAGTCT	TGCGCA	GTAGTATA	AGTCTT	900
DB										
QY	901	AAGAT	GATG	CAAA	ACCA	TGGCC	ATGAAA	TATC	ACCCT	960
DB										
QY	901	AAGAT	GATG	CAAA	ACCA	TGGCC	ATGAAA	TATC	ACCCT	960
DB										
QY	961	ATGGT	ATCT	CTCTT	TTTCAA	AGATG	GAA	CAATAG	GAGGAA	1020
DB										
QY	961	ATGGT	ATCT	CTCTT	TTTCAA	AGATG	GAA	CAATAG	GAGGAA	1020
DB										
QY	1021	GGTCT	CAATCT	CTTAGA	GATGAA	CA	CAGCAG	ACTGGT	GTATA	1080
DB										
QY	1021	GGTCT	CAATCT	CTTAGA	GATGAA	CA	CAGCAG	ACTGGT	GTATA	1080
DB										
QY	1081	ACCTT	AA	GTAG	CTTCC	CAGTGG	CA	GCTGG	ATGATG	1140
DB										
QY	1081	ACCTT	AA	GTAG	CTTCC	CAGTGG	CA	GCTGG	ATGATG	1140
DB										
QY	1141	GACC	AGCC	ATCT	ACT	GATGG	CT	CGATAT	ATGTTCT	1200
DB										
QY	1141	GACC	AGCC	ATCT	ACT	GATGG	CT	CGATAT	ATGTTCT	1200
DB										
QY	1201	ACTTT	GGG	ATTTG	CA	CCAT	TG	CCATG	ATGATG	1260
DB										
QY	1201	ACTTT	GGG	ATTTG	CA	CCAT	TG	CCATG	ATGATG	1260
DB										
QY	1261	GAA	GATAC	AGTCTT	GG	ACTT	GC	CA	CTG	1320
DB										
QY	1261	GAA	GATAC	AGTCTT	GG	ACTT	GC	CA	CTG	1320
DB										
QY	1321	ATG	ATTCT	AT	GATG	AGG	AGG	CA	TGATG	1380
DB										
QY	1321	ATG	ATTCT	AT	GATG	AGG	AGG	CA	TGATG	1380
DB										
QY	1381	ACATT	GCG	AGGA	CGCA	ATTGG	AGTCTT	CT	CGTGG	1440
DB										
QY	1381	ACATT	GCG	AGGA	CGCA	ATTGG	AGTCTT	CT	CGTGG	1440
DB										
QY	1441	AAATTT	CT	TAAG	CCGCT	CAAGCT	CT	CGCAT	GCTT	1500
DB										
QY	1441	AAATTT	CT	TAAG	CCGCT	CAAGCT	CT	CGCAT	GCTT	1500
DB										
QY	1501	TAC	AGTAT	CCT	GAG	AAATTT	GCC	AGG	AAATTTAT	1560
DB										
QY	1501	TAC	AGTAT	CCT	GAG	AAATTT	GCC	AGG	AAATTTAT	1560
DB										
QY	1561	CAGT	TCGG	AGAG	AA	CGC	AGCT	CT	GCAT	1620
DB										
QY	1561	CAGT	TCGG	AGAG	AA	CGC	AGCT	CT	GCAT	1620
DB										
QY	1621	CTGT	GTC	GCAT	CGTAT	TGGA	AGG	AAAT	TGAG	1680
DB										
QY	1621	CTGT	GTC	GCAT	CGTAT	TGGA	AGG	AAAT	TGAG	1680
DB										
QY	1681	ACAA	TTTGT	GG	CA	TG	CTGTG	TC	CGGG	1740

Db	1681	ACAATTTGTGGGCATGACATGTGGTCCGGGAGGACAGTGTGTAATATGTGTGAATGATGTGATGAA	1740
Qy	1741	GGCCCCAAGCCCAACCACTGGCCACTGGTCGCACTGCTCTTTGTGGTCCCATCTGCTCCAGG	1800
Db	1741	GGCCCCAAGCCCAACCACTGGCCACTGGTCGCACTGCTCTTTGTGGTCCCATCTGCTCCAGG	1800
Qy	1801	ACTGCGGAGGGGAGTATCTCATAGGAGTCCCTCTGCACCAACCCCAAGCCATCGCAT	1860
Db	1801	ACTGCGGAGGGGAGTATCTCATAGGAGTCCCTCTGCACCAACCCCAAGCCATCGCAT	1860
Qy	1861	GGAGGGAGTTCTGTGAGGGCTCCACTCGCACTCTGAAAGCTCTGCAACAGTCAGAAATGT	1920
Db	1861	GGAGGGAGTTCTGTGAGGGCTCCACTCGCACTCTGAAAGCTCTGCAACAGTCAGAAATGT	1920
Qy	1921	CCCCGGGACAGTGTGACTTCCGTGCTGCTCAGTGTGCCGAGCAACAGCAGACGATTC	1980
Db	1921	CCCCGGGACAGTGTGACTTCCGTGCTGCTCAGTGTGCCGAGCAACAGCAGACGATTC	1980
Qy	1981	AGAGGGCGGCACTACAGTGGAGCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAA	2040
Db	1981	AGAGGGCGGCACTACAGTGGAGCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAA	2040
Qy	2041	CTCTACTGTATCGCAGAAAGATTGATTTCTTTCTTTTGTCTCAATAAAGTCAAGAT	2100
Db	2041	CTCTACTGTATCGCAGAAAGATTGATTTCTTTCTTTTGTCTCAATAAAGTCAAGAT	2100
Qy	2101	GGGACTCCATGCTCGGAGGATAGCCGTAATGTTTGTAAGATGGGATATGTGAGAGATT	2160
Db	2101	GGGACTCCATGCTCGGAGGATAGCCGTAATGTTTGTAAGATGGGATATGTGAGAGATT	2160
Qy	2161	GGATGTGCAATGCTCTTGATCTGATCTGATGTTGAAGACGTCTGTGGGGTGTGTAACGGG	2220
Db	2161	GGATGTGCAATGCTCTTGATCTGATGTTGAAGACGTCTGTGGGGTGTGTAACGGG	2220
Qy	2221	AATACTCAGCCTGCGACGATTCACAGGGTCTCTACACCAAGCACCAACACCAACCCAG	2280
Db	2221	AATACTCAGCCTGCGACGATTCACAGGGTCTCTACACCAAGCACCAACACCAACCCAG	2280
Qy	2281	TATTATCAATGTGACCAATTCCTCTGGAGCCCGAGTATCCGATCTATGAATGAAC	2340
Db	2281	TATTATCAATGTGACCAATTCCTCTGGAGCCCGAGTATCCGATCTATGAATGAAC	2340
Qy	2341	GTCTTACCTCCTACATTTCTGTGCGAATGCCCTCAGAAAGTACTACCTGAATGGGAC	2400
Db	2341	GTCTTACCTCCTACATTTCTGTGCGAATGCCCTCAGAAAGTACTACCTGAATGGGAC	2400
Qy	2401	TGGAACGTGGACTGGCCCGCGGTACAAATTTTCGGGCACTACTTTGATACAGACGG	2460
Db	2401	TGGAACGTGGACTGGCCCGCGGTACAAATTTTCGGGCACTACTTTGATACAGACGG	2460
Qy	2461	TGCTTAATAGACCCGAGACTTAAATCGCTACTGGACCAACCAAGCAGACACTGATCTG	2520
Db	2461	TGCTTAATAGACCCGAGACTTAAATCGCTACTGGACCAACCAAGCAGACACTGATCTG	2520
Qy	2521	GAGCTGCTTTTCAGGGAAGGAAACCCGGGTGTGCTTGGGAATCTCCATCCCTGCTG	2580
Db	2521	GAGCTGCTTTTCAGGGAAGGAAACCCGGGTGTGCTTGGGAATCTCCATCCCTGCTG	2580
Qy	2581	GGGACCGGAGAGCAGCCCTGCCCCAGCCCAAGCTACACTTGGGCCATCTGGGCTCTGAG	2640
Db	2581	GGGACCGGAGAGCAGCCCTGCCCCAGCCCAAGCTACACTTGGGCCATCTGGGCTCTGAG	2640
Qy	2641	TGCTCCGTGCTCGGGAGGGGACAGATGACCGTGAGAGGGCTCTACAGAGACCTG	2700
Db	2641	TGCTCCGTGCTCGGGAGGGGACAGATGACCGTGAGAGGGCTCTACAGAGACCTG	2700
Qy	2701	AAGTTTCAAGTAAATATGTCTCTGCAATCCCAAGACACGCTGTCAAGGGCTGGT	2760
Db	2701	AAGTTTCAAGTAAATATGTCTCTGCAATCCCAAGACACGCTGTCAAGGGCTGGT	2760
Qy	2761	CCCTGCAAGTATCTGCTGCTCTCCAGCTGTCCGTGGGAACTGGAGTGGCTGCAGT	2820

2761 CCTTGCAGAGTATCTGCTGCTCTCCAGAGTGTGCTGGGGAACTGGAGTGCTGCTGAGT 2820
2821 CGGACGTGTGGCGGGGTGCGCAGAGCGCCCGTGCAGTGCAACAGCGGGGTGCACTAT 2880
2821 CGGACGTGTGGCGGGGTGCGCAGAGCGCCCGTGCAGTGCAACAGCGGGGTGCACTAT 2880
2881 GACTCGGAGCAGTCCGGGCGAGCCTGTGGCCCTCAGCCTGCTCTCCAGCAGGAGGCC 2940
2881 GACTCGGAGCAGTCCGGGCGAGCCTGTGGCCCTCAGCCTGCTCTCCAGCAGGAGGCC 2940
2941 TGCACCTCTCAGAGCTGCCACCTGTCATGAGAGCGCGCGCCCTGGGCGAGAGTCTCACAC 3000
2941 TGCACCTCTCAGAGCTGCCACCTGTCATGAGAGCGCGCGCCCTGGGCGAGAGTCTCACAC 3000
3001 ACCTGTGGGAAGGGGTGGAGGAAGCGGCGAGTGGCCCTGTAAAGACCAACCCCTCGGCC 3060
3001 ACCTGTGGGAAGGGGTGGAGGAAGCGGCGAGTGGCCCTGTAAAGACCAACCCCTCGGCC 3060
3061 AGAGCGCAGTGTGCGCCGACCGCTGTGCACTCTCGAGGCCAAGCCCGAGGATGCATGAA 3120
3061 AGAGCGCAGTGTGCGCCGACCGCTGTGCACTCTCGAGGCCAAGCCCGAGGATGCATGAA 3120
3121 GCCTGTCTCTCAGCGCTGCCACCAAGCCCAAGAGCTGCAGTGGCTGTGCTCGCCTGG 3180
3121 GCCTGTCTCTCAGCGCTGCCACCAAGCCCAAGAGCTGCAGTGGCTGTGCTCGCCTGG 3180
3181 TCCAGTGTCTGTGATGTGAAAGAGGAAACACAGAAAGATTCTTAAATGTGCTGAA 3240
3181 TCCAGTGTCTGTGATGTGAAAGAGGAAACACAGAAAGATTCTTAAATGTGCTGAA 3240
3241 AAGTATGTTTCTGGAAGTATCGAGAGCTGGCCCTCAAAGAGTGCTCAATTTGCCGAAG 3300
3241 AAGTATGTTTCTGGAAGTATCGAGAGCTGGCCCTCAAAGAGTGCTCAATTTGCCGAAG 3300
3301 CCCAGCTGGAGCTGAAAGTGGCTGGCGCCGCTTCCATGCCCGGAGCCGCCCATTT 3360
3301 CCCAGCTGGAGCTGAAAGTGGCTGGCGCCGCTTCCATGCCCGGAGCCGCCCATTT 3360
3361 GCTGCTGGCGGACCTCGAGGGGAGCTGGTTTGGCTCACCTGGTCTCAGTGCAAGGCC 3420
3361 GCTGCTGGCGGACCTCGAGGGGAGCTGGTTTGGCTCACCTGGTCTCAGTGCAAGGCC 3420
3421 AGCTGTGGGGAGCGTTTCAGACGAGTCCGTGCAAGTGTGCTGGCTGGGGCGCGCGCC 3480
3421 AGCTGTGGGGAGCGTTTCAGACGAGTCCGTGCAAGTGTGCTGGCTGGGGCGCGCGCC 3480
3481 TCAGGCTGCTCTCCAGCAGAGCCTTGGCGCTCCCTGGCTGCAACACTCACTCTCTGC 3540
3481 TCAGGCTGCTCTCCAGCAGAGCCTTGGCGCTCCCTGGCTGCAACACTCACTCTCTGC 3540
3541 CCNATTGCAGAGAAAGATGCTTCTGCAAGAGCTACTTCCACTGGTGCTACCTGGTA 3600
3541 CCNATTGCAGAGAAAGATGCTTCTGCAAGAGCTACTTCCACTGGTGCTACCTGGTA 3600
3601 CCCAGCAGGGGATGTGACGCCCAAGTGTACGGCAAGCAGTGTGCAAGTGTGCTCT 3660
3601 CCCAGCAGGGGATGTGACGCCCAAGTGTACGGCAAGCAGTGTGCAAGTGTGCTCT 3660
3661 AAGTCCAACTTGTGA 3675
3661 AAGTCCAACTTGTGA 3675

RESULT 3
US-09-930-872-5
; Sequence 5, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. 6448388el Human Proteases and Polynucleotides Encoding the Sa
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/09/930, 872

; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4042
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-930-872-5

Query Match 100.0%; Score 3675; DB 3; Length 4042;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3675; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGCCCCCGCGCGCGATGGCGGGCTTGGCGGCGCTGTGGATGCTGTCTGGGCGCAG 60
DB 99 ATGAAGCCCCCGCGCGCGATGGCGGGCTTGGCGGCGCTGTGGATGCTGTCTGGGCGCAG 158
QY 61 GTGGCGGAGCAGCAGCTGCGCCATGCGGACCCGCGAGCGGCGAGCGGCTGGAGCCCG 120
DB 159 GTGGCGGAGCAGCAGCTGCGCCATGCGGACCCGCGAGCGGCGAGCGGCTGGAGCCCG 218
QY 121 AGCGTCCCCGCTCTCTCCACCCGCGAGCGGCGGCGCTGGATGCAAGGCGGCAATAT 180
DB 219 AGCGTCCCCGCTCTCTCCACCCGCGAGCGGCGGCGCTGGATGCAAGGCGGCAATAT 278
QY 181 GACCTGTGCTCTGCGCTTACGAGGTTGACACAGGGGCGATTAACGTGTCCTCAATGAATCATG 240
DB 279 GACCTGTGCTCTGCGCTTACGAGGTTGACACAGGGGCGATTAACGTGTCCTCAATGAATCATG 338
QY 241 CACCATACGCGCGGAGAGAGCAGTGGCGGTCCGAGGTTGAGTCTCTTCACTTCGCG 300
DB 339 CACCATACGCGCGGAGAGAGCAGTGGCGGTCCGAGGTTGAGTCTCTTCACTTCGCG 398
QY 301 CTGAAAGGCTTCCAGGACGACTTCCAGCTGGATCTGAGGACTTCCAGGACCTAGTGGCT 360
DB 399 CTGAAAGGCTTCCAGGACGACTTCCAGCTGGATCTGAGGACTTCCAGGACCTAGTGGCT 458
QY 361 CTGCGGCTTTATGTGTCAGACGTTGGGAAAGACAGGCACTAAAGTCTGTGACAGCTTTACCG 420
DB 459 CTGCGGCTTTATGTGTCAGACGTTGGGAAAGACAGGCACTAAAGTCTGTGACAGCTTTACCG 518
QY 421 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGCGATCAACAGAACTCTCAGTGGCC 480
DB 519 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGCGATCAACAGAACTCTCAGTGGCC 578
QY 481 CTTTCACTTCCAGGCTTGTTCAGGCTATGATACGAAACAGAGAGGCGAGTTACTTCTTA 540
DB 579 CTTTCACTTCCAGGCTTGTTCAGGCTATGATACGAAACAGAGAGGCGAGTTACTTCTTA 638
QY 541 AGGCGCACTTCTTCAACCTCTCATGGAACCTCGGAGAGCTGCCCAAGGCGAGCTCGCCA 600
DB 639 AGGCGCACTTCTTCAACCTCTCATGGAACCTCGGAGAGCTGCCCAAGGCGAGCTCGCCA 698
QY 601 TCCACAGTACTGTATCAAGAGATCCACAGAGCCCAATGCTCTGGGGCGCAGTGAAGTCTG 660
DB 699 TCCACAGTACTGTATCAAGAGATCCACAGAGCCCAATGCTCTGGGGCGCAGTGAAGTCTG 758
QY 661 GTGACCTCAAGGACATGGGAGCTGGGCACTCAACCCCTGCAAGGAGGAGCTTCCGCTG 720
DB 759 GTGACCTCAAGGACATGGGAGCTGGGCACTCAACCCCTGCAAGGAGGAGCTTCCGCTG 818
QY 721 GGACTGCCACAAAAGCAGCAATTTCTGTGAAGACGCAAGAAATACATGCCCCAGCTCCC 780
DB 819 GGACTGCCACAAAAGCAGCAATTTCTGTGAAGACGCAAGAAATACATGCCCCAGCTCCC 878
QY 781 AAGGAAGACTCTTCTCATCTTGGCAGATGATTAAGTCTTGTGTACGGCATAGCGCTCT 840
DB 879 AAGGAAGACTCTTCTCATCTTGGCAGATGATTAAGTCTTGTGTACGGCATAGCGCTCT 938
QY 841 CTTTCTGAGTCTCCATAGAAATGAAGAACTGAAGCTGGAGACCTTGTGTGTGTCGACAAA 900

Db 1179 ACCTTAAGTAGCTTCTCCAGTGGCAGCTGTGGATTGATGGGAAAGATGGGACTGGTCAAT 1238
Qy 1141 GACCACGCCACTTACTAGACTGGTCTGGGATATATGTTCTTGGGAATAGGCCCTGTGCAC 1200
Db 1239 GACCACGCCACTTACTAGACTGGTCTGGGATATATGTTCTTGGGAATAGGCCCTGTGCAC 1298
Qy 1201 ACTTTGGGATTTGCACCCATAAGTGGAAATGTGTAGTAAATATCGCAGCTGACACGATTAAAT 1260
Db 1299 ACTTTGGGATTTGCACCCATAAGTGGAAATGTGTAGTAAATATCGCAGCTGACACGATTAAAT 1358
Qy 1261 GAAGATACAGGCTCTTGGACTGGCCTTCACCATTTGCCATGAGTCTGGACACAACTTTGGC 1320
Db 1359 GAAGATACAGGCTCTTGGACTGGCCTTCACCATTTGCCATGAGTCTGGACACAACTTTGGC 1418
Qy 1321 ATGATTTCAATGATGAGAGGGAAACATGTGTAAAGTCCGAGGGCAACATCATGTGCCCT 1380
Db 1419 ATGATTTCAATGATGAGAGGGAAACATGTGTAAAGTCCGAGGGCAACATCATGTGCCCT 1478
Qy 1381 ACATTTGGCAGGACGCAATGGAGTCTTCTCGGTCAACCTCGAGCCGCGAGTATCTACAC 1440
Db 1479 ACATTTGGCAGGACGCAATGGAGTCTTCTCGGTCAACCTCGAGCCGCGAGTATCTACAC 1538
Qy 1441 AAATTTCTAAGCACCGCTCAAGCTATCTGCCTTGTGATCAGCCAAAGCCTGTGAAGGAA 1500
Db 1539 AAATTTCTAAGCACCGCTCAAGCTATCTGCCTTGTGATCAGCCAAAGCCTGTGAAGGAA 1598
Qy 1501 TACAAGTATCTTGAGAAATTCGAGGAGAAATATATGATGATCAGCCAAAGCCTGTGAAGGAA 1560
Db 1599 TACAAGTATCTTGAGAAATTCGAGGAGAAATATATGATGATCAGCCAAAGCCTGTGAAGGAA 1658
Qy 1561 CAGTTCCGAGAGAAAGCAAGCTCTGCACTGCGACTTTAAAGAGGACATCTCTAAGCC 1620
Db 1659 CAGTTCCGAGAGAAAGCAAGCTCTGCACTGCGACTTTAAAGAGGACATCTCTAAGCC 1718
Qy 1621 CTGTGGTCCCATCTGATTTGGAGGAAATGTGACACTAAATTTATGCAAGCAGTGTGAAGCC 1680
Db 1719 CTGTGGTCCCATCTGATTTGGAGGAAATGTGACACTAAATTTATGCAAGCAGTGTGAAGCC 1778
Qy 1681 ACAATTTGTGGGCATGATGTGGTCCGGGGAGGACAGTGTGAAATATGATGATGAA 1838
Db 1779 ACAATTTGTGGGCATGATGTGGTCCGGGGAGGACAGTGTGAAATATGATGATGAA 1800
Qy 1741 GGGCCCAAGCCCAACCATGGCCACTGGTGGACTGTCTTGTGTCCTCCCATGTCCAGG 1860
Db 1839 GGGCCCAAGCCCAACCATGGCCACTGGTGGACTGTCTTGTGTCCTCCCATGTCCAGG 1898
Qy 1801 ACCTGGGAGGGGAGATATCTCATAGGAGTCGCTCTGACCAACCCCAAGCCATCGCAT 1860
Db 1899 ACCTGGGAGGGGAGATATCTCATAGGAGTCGCTCTGACCAACCCCAAGCCATCGCAT 1958
Qy 1861 GGAGGGAAGTTCTGTGAGGGCTCCACTGCACTCTGAAGCTCTGCAACAGTCAAGAAATGT 1920
Db 1959 GGAGGGAAGTTCTGTGAGGGCTCCACTGCACTCTGAAGCTCTGCAACAGTCAAGAAATGT 2018
Qy 1921 CCCGGGACAGTGTGACTTCGGTGTGCTCAGTGTGCGGAGCACAACAGCAGACGATTC 1980
Db 2019 CCCGGGACAGTGTGACTTCGGTGTGCTCAGTGTGCGGAGCACAACAGCAGACGATTC 2078
Qy 1981 AGAGGGCGGCATACAAAGTGGAGCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAA 2040
Db 2079 AGAGGGCGGCATACAAAGTGGAGCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAA 2138
Qy 2041 CTCTACTGTATCGCAGAGGATTTGATTTCTTTCTTTCTTTGTCAAATTAAGTCAAGAT 2100
Db 2139 CTCTACTGTATCGCAGAGGATTTGATTTCTTTCTTTCTTTGTCAAATTAAGTCAAGAT 2198
Qy 2101 GGGACTCCATCTCGGAGGATAGCGGTAAATGTTTGTATAGATGGGATATGTGAGAGATT 2160
Db 2199 GGGACTCCATCTCGGAGGATAGCGGTAAATGTTTGTATAGATGGGATATGTGAGAGATT 2258
Qy 2161 GGATGTGCAATCTCTTGGATCTGATGCTGTGGAAGCCTCTGCGGGTGTATCGGG 2220

Db 2259 GGATGTGCAAAATGTCTTGGATCTGATCTGTTGAAGACGTCTGTGGGGTGTGTAAACGGG 2318
Qy 2221 AATAACTCAGCCTGACGATTTACAGGGGTCTCTACCAAGCACCAACCAACCAACCAAG 2280
Db 2319 AATAACTCAGCCTGACGATTTACAGGGGTCTCTACCAAGCACCAACCAACCAACCAAG 2378
Qy 2281 TATTATCACATGTGTACCAATCTCTTCTGGAGCCGAGATATCCGATCTATGAATGAAC 2340
Db 2379 TATTATCACATGTGTACCAATCTCTTCTGGAGCCGAGATATCCGATCTATGAATGAAC 2438
Qy 2341 GTCTCTACCTCTCTACATTTCTGTGTGGCAATGCCCCCTCAGAAAGTACTACCTGAATGGGAC 2400
Db 2439 GTCTCTACCTCTCTACATTTCTGTGTGGCAATGCCCCCTCAGAAAGTACTACCTGAATGGGAC 2498
Qy 2401 TGGACCGTGGACTGGCCCGGGCGGTACAAATTTTGGGGCACTACTTTTCGACTACAGAGCG 2460
Db 2499 TGGACCGTGGACTGGCCCGGGCGGTACAAATTTTGGGGCACTACTTTTCGACTACAGAGCG 2558
Qy 2461 TCTTATAATGAGCCGAGAACTTAAATCGCTACTGGACCAACCAACGAGACACTGATTGTG 2520
Db 2559 TCTTATAATGAGCCGAGAACTTAAATCGCTACTGGACCAACCAACGAGACACTGATTGTG 2618
Qy 2521 GAGCTGTCTTTCAGGGAAGGAAACCCGGGTGTTCGCTTGGGAATACTCCATGCTCGCTTG 2580
Db 2619 GAGCTGTCTTTCAGGGAAGGAAACCCGGGTGTTCGCTTGGGAATACTCCATGCTCGCTTG 2678
Qy 2581 GGGACCGAGAGACGCCCTCGCCAGCCAGCTACACTTTGGGCCATCTGTGCGCTCTGAG 2640
Db 2679 GGGACCGAGAGACGCCCTCGCCAGCCAGCTACACTTTGGGCCATCTGTGCGCTCTGAG 2738
Qy 2641 TGCTCCGTCTCTGCGAGGGGACAGATGACCGTGGAGAGGGCTGTCTACAGAGACCTG 2700
Db 2739 TGCTCCGTCTCTGCGAGGGGACAGATGACCGTGGAGAGGGCTGTCTACAGAGACCTG 2798
Qy 2701 AAGTTTCAAGTAAATATGTCTTCTGGAATCCCAAGACACGACCTGTACGGGGCTGGTG 2760
Db 2799 AAGTTTCAAGTAAATATGTCTTCTGCAATCCCAAGACACGACCTGTACGGGGCTGGTG 2858
Qy 2761 CCTTGCAAGTATCTGCTGTCTCCAGCTGGTCCGTGGGGAATGGAAGTCCCTGCAAT 2820
Db 2859 CCTTGCAAGTATCTGCTGTCTCCAGCTGGTCCGTGGGGAATGGAAGTCCCTGCAAT 2918
Qy 2821 CGACAGTGTGGGGGGGTGCCAGAGCCGCCGTGTGCAAGTGCACAGCGGGGTGCATAT 2880
Db 2919 CGACAGTGTGGGGGGGTGCCAGAGCCGCCGTGTGCAAGTGCACAGCGGGGTGCATAT 2978
Qy 2881 GACTCGGAGCCAGTCCCGGCCAGCCTGTGCGCTCAGCTGTGCGGCGCCCTGGGAGAGTGTCAAC 2940
Db 2979 GACTCGGAGCCAGTCCCGGCCAGCCTGTGCGCTCAGCTGTGCGGCGCCCTGGGAGAGTGTCAAC 3038
Qy 2941 TGCACCTCTCAGAGTGTGCCACCTGCAATGAGAGCGCCGGCCCTGGGAGAGTGTCAAC 3000
Db 3039 TGCACCTCTCAGAGTGTGCCACCTGCAATGAGAGCGCCGGCCCTGGGAGAGTGTCAAC 3098
Qy 3001 ACTGTGGGAAGGGGTGGAGGAAGCGGCACTGTGCGCTGTAGAGACCAACCCCTCGGCC 3060
Db 3099 ACTGTGGGAAGGGGTGGAGGAAGCGGCACTGTGCGCTGTAGAGACCAACCCCTCGGCC 3158
Qy 3061 AGAGCGCAGCTGTCTCCCGACGCTGTGTGCACTCCGAGCCCAAGCCAGGATGATGAA 3120
Db 3159 AGAGCGCAGCTGTCTCCCGACGCTGTGTGCACTCCGAGCCCAAGCCAGGATGATGAA 3218
Qy 3121 GCCTGTCTGCTTTCAGCGCTGCCCAAGCTGTGTGCACTCCGAGCCCAAGCCAGGATGATGAA 3180
Db 3219 GCCTGTCTGCTTTCAGCGCTGCCCAAGCTGTGTGCACTCCGAGCCCAAGCCAGGATGATGAA 3278
Qy 3181 TCCAGTGTCTGTGAATGTGAAGAGAACACAGAAAAAGATTCTTAAATGTGCTGAA 3240
Db 3279 TCCAGTGTCTGTGAATGTGAAGAGAACACAGAAAAAGATTCTTAAATGTGCTGAA 3338
Qy 3241 AAGTATGTCTTGGAAAGTATCGAGAGCTGGCCCTCAAGAGTGTCTCAATTGGCCGAAG 3300
Db 3339 AAGTATGTCTTGGAAAGTATCGAGAGCTGGCCCTCAAGAGTGTCTCAATTGGCCGAAG 3398

;; TITLE OF INVENTION: PROTEASE HOMOLOGS
;; FILE REFERENCE: 5800-55
;; CURRENT APPLICATION NUMBER: US/09/392,184
;; CURRENT FILING DATE: 1999-09-09
;; NUMBER OF SEQ ID NOS: 33
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 5
;; LENGTH: 5357
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)...(5357)
;; OTHER INFORMATION: reprotolysin (ADAM family of metalloprotease)
;; NAME/KEY: misc feature
;; LOCATION: (1)...(5357)
;; OTHER INFORMATION: n = A,T,C or G
US-09-392-184-5

Query Match 8.3%; Score 306.2; DB 3; Length 5357;
Best Local Similarity 51.3%; Pred. No. 1.4e-81;
Matches 1123; Conservative 0; Mismatches 1003; Indels 65; Gaps 15;

QY	863	AGAACTGAACGTGGAGACCTTGGTGTGTCACAAAAGATGATGCAAAACCATGGCC	922
DB	4600	AGAGAAGTGGGTGGAGACCTGGTAGTACTGATGTCGCAAAATGGTGGAGTACCAGGAC	4541
QY	923	ATGAAATATCACCACTAGCTCTCAGCATCTCAACATGATGATCTGCTTTATTCAAAG	982
DB	4540	AGCCGAGGTGAGAGCTATGTCTGACCATCATGAACATGATGGTGGCTGTTTCATG	4481
QY	983	ATGGAACAATAGAGAGAAACATCAACATTTGCAATGTAGTCTGATCTTTAGAAAGATG	1042
DB	4480	ACCCAGCATTTGGAAACCCCATCCATCAACATTTGGGCTGCTGCTGCTGGAAGATG	4421
QY	1043	ACAGCCAGAGCTGGTGAAGTACACGACGACACACACCTTAAGTAGCTTCGCCAGT	1102
DB	4420	AGGAGGAGGACCTTAAAGATCAGCAACATGACAGACCAACCCCTGAAGAGCTTCGCAAGT	4361
QY	1103	GGCAGTCTGATTTGATGGGGAAGATGGGACTCGT-----CATGCCACGCCATCT	1153
DB	4360	GGCAGAAAGCATCAACATGAAGGGGATGCCCATCCCTGCCACATGACATGCCATCC	4301
QY	1154	TACTGACTGGTCTGGATATATGTTCTTGAAGAAATGAGCCCTGTGTGACATTTGGGATTTG	1213
DB	4300	TGCTCACGAAAGGACCTGTGTGACGACCATGAACCGGCCCTGTGAGACCTGGGACTGT	4241
QY	1214	CACCAATAAGTGAATGTAGTAAATNTGCGAGCTGCGAGATTATGAAGATACAGGTC	1273
DB	4240	CCCATGTGGGGCATGTGCCAGCCGACCGCAGCTGCGAGCATCAACGAGGACACGGCC	4181
QY	1274	TTGGACTGGCTTCACCATTTGCCATGTGTCGACACAACTTTGGCATGATTCATGATG	1333
DB	4180	TGCCGTGGCCCTTCACTGTAGCCACAGACTGGGACACAGTTTGGCATTCAGATGACG	4121
QY	1334	GAGAAGGGAACATGTGTAAAGTCCGAG-----GGCAACATCATGTCCCTPACATGG	1387
DB	4120	GAAGCGCAATGACTGTGAGCCGTTGGGAAACGACCTTTTCATCATGTCTCCACAGCTCC	4061
QY	1388	CAGGACGCAATGGAGTCTTCTCTGTGTCACCTCTGAGCGCCGAGTATCTACAAATTTTC	1447
DB	4060	TGTACGACCGCTCCCTCCCTCACCCTGTGTCGCTGCGAGCGCCAGTATATCACAGGTTCC	4001
QY	1448	TAAAGCACCGCTCAAGCTATCTGCTTGTGATCAGGCCAAAGCCTGTGAAGGAATACAAAGT	1507
DB	4000	TTGACCGTGGGTGGGCTGTGCTGACGACCTCC---TGCAAGGACATTTATCGACT	3944
QY	1508	ATCCTGAGAAATTTGCCAGGAAATATATGATGCAAAACACACAGTGCAGTGGCAGTTCC	1567
DB	3943	TCCCTCGGTGCGCACTGGCGTCTCTATGATGAAGCCACCAAGTGGCGCTCCAGTACG	3884
QY	1568	GAGAGAAAGCAAGCTCTGATGCTGCACTTTTAAAGGACATCTGTAAAGCCCTGTGGT	1627

DB	3883	GGGCTACTCTGCTTCTGC-----GAGGACATGATATGTCTGCCACACACTCTGGT	3830
QY	1628	GCCATGATATGGAAGAAATGTGAGACTAAATTTTATGCCAGCAGAGGACAAATTT	1687
DB	3829	GC--TCTGTGGGACCACTGTCACTTCAAGCTGGATGAGCCGCTGGAGCCGACCCGGT	3772
QY	1688	GTGGGATCATATGTGTGCTGGGAGGACAGTGTGTAATATGTGTGATGAAGGCCCA	1747
DB	3771	GTGGGGAGAAATAAGTGTGTCTCAGTGGGAGTGGCT---ACCCGTGGGCTTCCGCCCG	3715
QY	1748	AGCCACCATGCGCACTGCTCGGACTGCTTCTTGTGCTCCCATCTCTCAGGACCTGCG	1807
DB	3714	AGCCGTGATGTGTGCTGTGCTGTGCTGAGCGCTGTGCTCATCTCTCAGGAGCTGTG	3655
QY	1808	GAGGGGAGTATCTCATAGAGTGTGCTCTGCAACCAACCCCAAGCCATCGCATGAGGGA	1867
DB	3654	GCATGGGCGTACAGAGCGCGGAGCGGAGTGCAGCGAGCTAGCGCCAAATACAAAGCA	3595
QY	1868	AGTTCTGTGAGGGCTCCACTCGCACTCTGAAGCTCTGCAACAGTCAAGAAATGTCCCGG	1927
DB	3594	GATACTGTGTGGGTGAGCGCAAGCGCTTCGGCTCTGCAACCTGCGAGGCTGCCCTGTG	3535
QY	1928	ACAGTGTTCACATTCGGTGTGCTCAGTGTGCGGAGCACAACAGCAGACGATTCAGAGGGC	1987
DB	3534	CGCGCCCTCTTCCGCCAGTCCAGTGCAGCACTTTGACGCTATGCTCTACAAGGGCC	3475
QY	1988	GGCCTACAAGTGGAGCCTTACACTCAAGTGAAGATCAGGACTTTATGCAACTCTACT	2047
DB	3474	AGCTGCACACATGGGTGC-----CGTGTCAATGACGTGAACCCCTGCGAGCTGCAC	3421
QY	2048	GTATCGCAGAGGATTTGATTTCTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2107
DB	3420	GCCGCCCCGCAATGATGACTTTTGGCGAGAGCTGCGGAGCGCGTGGTGCATGGCACCC	3361
QY	2108	CATGCTCGAGGA-----TAGCCGTAATGTTTGTATAGATGGGATATGTGAGAGAGTTG	2161
DB	3360	CTGTCTACAGTCCGAGCGGACCTCTGCAATCAACGGCATCTGTGAAGAGTGG	3301
QY	2162	GATGTACAATGTCTTGGATCTGATGTGTTGAAGACGTCTGTGGGGTGTGTAACGGGA	2221
DB	3300	GCTGTGACTTCGAGATTGACTCCGGTGTCTATGAGGACCGCTGTGTGTGTGCCACGGCA	3241
QY	2222	ATAACTCAGCTGCAGGATTCACAGGGGTCTCTACACCAAGCCACCAACCAACCACT	2281
DB	3240	ACGGCTCCACCTGCCACACCGTGAAGCGGACCT---TCGAGGAGCGCGGCTCGGGT	3184
QY	2282	ATTATCACAATGTCAACATTTCTTCTGAGCGCGGAGTATCCGCATCTATGAATGAACG	2341
DB	3183	ATGTGATGTGGGTGTATCCAGCGCGCAGCGGAGATCCGCATCCAGAGGTGGCG	3124
QY	2342	TCTCTACCTCTCACAATTTCTGTGCGCA---ATGCCCCCTCAGAAGGTACTTACCTGAATGGG	2398
DB	3123	AGGCTGCCAACTTCTTGGCACTGCGGAGTGAGGACCGGAGAGTACTTCTCAATGGTG	3064
QY	2399	ACTGAGCCGTGACTGCGCGCGGTACAAATTTTCGGGGCACTACTTTGAGTACAGAC	2458
DB	3063	GCTGGACCAATCCAGTGGAAACGGGAGCTACAGGTGGCAGGAGCCACCTTCACATACGCAC	3004
QY	2459	GGTCTCTAATAGCCCGGAGAACTTAAATGCTTCTGAGCAACCAACAGAGACACTCATTTG	2518
DB	3003	GAGGGGCAACTGG---GAGAACCTCAGCTCCCGGGTCCACCAAGAGGCTGTCTGGA	2947
QY	2519	TGGAGTGTCTGTTTTCAGGGAAGAAACCGGGTGTGCTTGGGAATACTCATGCTCGCT	2578
DB	2946	TCCAGCTGTCTTCCAGGAGAGCAACCTTGGGTGCACTACGAGTACACCATCCACAGGG	2887
QY	2579	TGGGAGCCGAGAGAGCGCCCTGCCAGCCAGCTACACTTGGGGCATCTGTCGCT---	2635
DB	2886	AGGCAAGTGGCCACAGCAGGCTCCCGCCCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2827
QY	2636	CTGAGTGTCTCTGCTTCTGCGGAGGGGACAGTGAACCTGTGAGAGAGGCTGTCTACAG	2695
DB	2826	CCAAGTGCACAGTCACTTCCGCGAGAGGTGTGAGAGGCGAGATGTGTACTTCTTGAGC	2767

Db 2098 TTCTGAATTTGACAGCATCCCTTTTCGTGGAAATTTCTACAAGTGGAAACGTACCGGG 2157
QY 2016 AGTAGAAGATCAGGACTTATGCAAACTCTACTGTATCGCAAGAGATTTGATTTCTTCTT 2075
Db 2158 AGGGGGGTGAAGGCC---TGCTCGCTCAGTGCCTAGCGAAGGCTTCACTTCTACAC 2214
QY 2076 TTCTTTGTCAATAAAGTCAAGATGGGACTCCATGCTCGGAGATAGCGTAAATGTTG 2135
Db 2215 GGAGAGGGCGGACGCGTGTGGACGGGACACCTTCCGCTCCAGACACGCTGGACATTTG 2274
QY 2136 TATAGATGGGATATGAGAGATTTGATGTGCAATGCTTGGATCTGATCTGTGTA 2195
Db 2275 CGTCAGTGGGGAATGCAAGCAGCTGGGCTCGGACCGAGTCTTGGGCTCGACCTCGGGGA 2334
QY 2196 AGACGCTGTGGGGTGTGTAACGGGAATTAACCTGAGCTGCACGATTCACAGGGGCTCTCTA 2255
Db 2335 GGACAAGTGCAGTGTGTGGCGGTGACGCGAGTGCCTGCGAGACCATCGAGGGCTCTT 2394
QY 2256 CACCAAGCACACACACCAACAGTATTAATCATGGTCAACATTCCTTCTCGGAGCCG 2315
Db 2395 CAGCCAGCCTCACTCGGGCGGGTACGAGGATGCTGTGATTCACCAAGGCTCCGT 2454
QY 2316 GAGTATCCGATCTATGAATGAAGTCTCTACTCTCTACATTTCTGTGGCAATGCCCT 2375
Db 2455 CCACATCTTCTATCCAGATCTGAACCTCTCTCTCAGTCACTTGGCCCTGAAGGGAGACCA 2514
QY 2376 CAGAAGGTACTACTGAATGGGCACTGGACCGTGGACTGGCCCGCGGTACAAATTTTC 2435
Db 2515 GGAGTCCCTGCTGTGAGGGGTGCTGCTGGGACCCCGACCCACCGTCTGCTTAGC 2574
QY 2436 GGGCACTACTTTTCGACTACAGACGGTCTATAATAGCCGAGAACTTAATCCTACTGG 2495
Db 2575 TGGGACCACTTTCACTGCGACAGCGGGCCAGACAGGTCACAGGCTCGAAGCCCTGG 2634
QY 2496 ACCAACCAAGACACACTGATTTGTGAGTCTGTTTTCAGGGAAGAAACCGGGTGTGC 2555
Db 2635 ACCGATTAATGCAATCTCTCATGCTCATGCTGTGCTGGGCC---GGACCGAGTCTGCTCCCT 2691
QY 2556 CTGGGAATACTCCTGCTGCTTGGGACCGAGAGACAGCCCTGCTGCCAGCCAGCTA 2615
Db 2692 CCGTACCGCTTCAATGCTCCCAATCGCCGCTGACTGCTGCCCCCTACTCTCTGGCACTA 2751
QY 2616 CACTTGGGCACTCGTGGCTCTGAGTCTCGCTGCTCGGAGGGGAGCAGATGACCGT 2675
Db 2752 TGGCCCTGACCAAGTCTCGG---CCGAGTGTGACGGCGGTAGCCAGGTGCGAGC 2805
QY 2676 GAGAGGGTCTGACAGACCTGAAGTTTCAAGTAAATATGCTCTCTGCAATCCCA 2735
Db 2806 GGTGGAGTGGCGCAACAGCTGGACAGCTCGCGGTGCGCCCCCACTACTGCAAGTGC 2865
QY 2736 GACACGACCTGTACCGGGGTGTGCTTCAAGATATCTGCTGTCTCCAGCTGTGTC 2795
Db 2866 CAGCAAGCTGCCCAAGAGGACGCGCTGCAACACGAGGCTTGGCCCTCCAGACTGGGT 2925
QY 2796 CGTGGGGAATGAGTGTCTGCAAGTGGAGTGGGGGTGCCAGAGCGCCCGCT 2855
Db 2926 TGTAGGGAATGCTGCTGTGAGCCGCGAGCTCGATGCAAGGTGCGAGTGTGCTGGT 2985
QY 2856 GCATGACACAGCGGGGT---GCATATGATCTCGAGCCAGTCCCGCCAGCTGTGCC 2912
Db 2986 CGTGTGCGCGCGCGGTCTCTCGCGGAGGAGAGGGGCTGGAACAGCGCATGCGCC 3045
QY 2913 TCAGCTGTCTCCCTCAGCAGGAGCGCTGCACTCTCAGAGTGCCTCCAGCTGAG 2972
Db 3046 GCAGCGCGCCCACTTGTACTGAGGCTGCGCACGCGCCCACTTGGCCCTCCGAGTGG 3105
QY 2973 CGCGGGCCCTGGGAGAGTGTCTACACACTGTGGAGAGGGGTGGAGGAGCGGCGAGT 3032
Db 3106 GGCCTCTGACTGTCTGAGTGAACCCCGAGCTCGGGCGGGGCTCGGCCACCGCGTGGT 3165
QY 3033 GGCTGTGAAGACCAACCCCTCGGCCAGAGCGCAGCTGTGCTGCCGACGCTGTGTCAC 3092

Db 3166 CTTTGCAGAGCGCAGACCCAGCGCCACGCTGCCCCCGGCGCACTGTCTACCCGCGCG 3225
QY 3093 C 3093
Db 3226 C 3226
RESULT 9
US-09-981-953A-3
; Sequence 3, Application US/09981953A
; Patent No. 6689599
; GENERAL INFORMATION:
; APPLICANT: RACIE, LISA A.
; APPLICANT: TWINE, NATALIE C.
; APPLICANT: AGOSTINO, MICHAEL J.
; APPLICANT: WOLFMAN, NEIL
; APPLICANT: MORRIS, ELISABETH A.
; TITLE OF INVENTION: NOVEL AGGRECANASE MOLECULES
; FILE REFERENCE: 08702.0075-00000
; CURRENT APPLICATION NUMBER: US/09/981,953A
; PRIOR FILING DATE: 2001-10-18
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3377
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Nucleotide
; OTHER INFORMATION: sequence of the aggrecanase molecule
US-09-981-953A-3
Query Match 8.0%; Score 295.4; DB 4; Length 3377;
Best Local Similarity 48.9%; Pred. No. 2e-78;
Matches 1115; Conservative 0; Mismatches 1091; Indels 75; Gaps 9;
QY 864 AGAACTGAACGCTGGAGACCTTGGTGGTTCGACAAAAGATGATGCAAAACCATGGCCA 923
Db 722 AGAGCGCTAGCTGGAGACCTTGGTGGTTCGACAGATGATGTTGGCTTATCAGGGCG 781
QY 924 TGAANAATACACACCTACGCTCAAGATCTCAACATGATGATCTGCTTTATTCAAAGA 983
Db 782 CCGGGATGTGGAGCAGTATGCTCGCCCATCATGAACATTTGTGCCAAACTTTTCCAGGA 841
QY 984 TGGACAATAGGAGGAAACATCAACATTTGCAATTTAGTCTGATTTCTTCTAGAAATGA 1043
Db 842 CTGAGTCTGGGAGACCGCTTAAACATCTCGTAACTCGCCTCATCTGCTCAGGGGA 901
QY 1044 ACAGCCAGGACTGGTGAATAGTACCAACGACAGACCAACCTTAAAGTAGCTTCTGCCAGTG 1103
Db 902 CAGGCCACTCTGGAGTACCCACCATCGCGGGAAGTCCCTGGACAGCTTCTGTAGTG 961
QY 1104 GCAG-----TCTGATTTGATCGGGAAGATGGAC 1133
Db 962 GCAGAAATCCATCGTGAACACACAGCGGTCAATGCAATGCCATTCACAGAACGGTGGC 1021
QY 1134 TGTCTATGACAGCCCATCTTACTGCTGTCTGGATATATGTTCTCTGGAAGATGAGCC 1193
Db 1022 TAAACATGACACAGCAGTGTCTCATCACGCTATGACATCTGCATCTCAAGAACAACCC 1081
QY 1194 CTGTGACACTTTGGGATTTGGACCCCAATAGTGAATGTGTAGTAAATATTCGAGCTGCAC 1253
Db 1082 CTGCGGCACTATAGGCTGTGCCCCCGGTGGCGGAATGTGTAGCGCGGAGAAAGCTTCAG 1141
QY 1254 GATTAATGAAGATACAGGCTTTGGAGTGGCCTTTCACATTGGCCCATGAGTCTTGACACA 1313
Db 1142 CGCTAATGAGGACATTTGGCCTGSCACACAGCGTTTCACTATTGCCACGAGATCGGCACAC 1201
QY 1314 CTTTGGCATGATTCATGATGAGAGAGGAAACATGT-----GTAAAAAGTCCGAGG 1364
Db 1202 ATTCGCGCATGAACCATGACGGGTGGGAAACAGCTGTGGGGCGCGTGGTTCAGGACCCAGC 1261

1365 CAACATCATGTCCCTACATTTGGCAGAGCGCATGGAATCTTCTCCTGGTCAACCTGCGAG 1424
1262 CAAGCTCATGGCTGCCCAATTAACATGAAGACCAACCCGTTGCTGTGGTCACTCGCAG 1321
1425 CCGCCAGTATCTACACAATTTCTAAGCACCGCTCAAGCTATCTGCGCTTGTGATCAGCC 1484
1322 CCGTGACTACATCACCAGCTTCTAGACTCGGGCTGGGGCTGTGCGCT---GAACAACCG 1378
1485 AAAGCCTGTGAAGGAATACAGATATCTTGAGAAATGCCAGGAGAATATATATGACNA 1544
1379 ACCCCCGACAGAGGACTTTGTGTACCCGACAGTGCACCGGGCCAGGCTACGATGAGA 1438
1545 CACACAGTGCAGAGTGGCAGTTCTGAGAGAGAGACCAAGCTCTGCAATGCTGACATTTAAAA 1604
1439 TGAGCAATGCCGCTTTCAGCATGAGTCAATCCGCTCAGTGTGTA-----AATACGG 1489
1605 GGACATCTGTAAAGCCCTGTGGTGCATCTGTAATTTGGAAGGAATGTGAGACTAAATTTAT 1664
1490 GGAAGTCTGCAGCGAGCTGTGGTGTCTGAGCAAGAGCAACCGGTGCATCACCACAGCAT 1549
1665 GCCAGCAGCAGAGGACAAATTTCTGGG-----CATGACATGTGTGGCGGGAGG 1715
1550 CCGCGCGCCGAGGGCAGCTGTGCCAGACGCAACCAATCGACAGGGGTGTGTCTACAA 1609
1716 ACAGTGTGTAAATATGGTGATGAAGGCCCCCAAGCCCAACCATGGCCACTGTGTCGACTG 1775
1610 ACGGGTCTGTGCCCTTTGGGTGGCGCCAGAGGGTGTGGAAGGAGGCTGTGCTCTTCTAGCGTCA 1729
1776 GTCTTCTTGTGCCCATCTCCAGGACCTCGGAGGGGGAGTATCTOATAGGAGTGCCT 1835
1670 GACTCCATGGGCGAGCTGACAGCGGACCTGTGGCGGGCGGTGTCTCTTCTAGCGTCA 1729
1836 CTGCACCAACCCCAAGCCATCGCATGGAGGGAAGTTCTGTGAGGGCTCCACTCGCACTT 1895
1730 CTGCGACAGCCCGAGCCCAACCATCGGGGCAAGTACTGTCTGGGTGAGAGAGCGGCA 1789
1896 GAAGCTCTGCAACAGTCAAGAAATGTCCCGGGACAGTGTGACTTCCGTGCTGCTCAGTG 1955
1790 CCGCTCTGCAACAGCAAGATGACTGTCCCTGCTCCAGGACTTCAGAGAGTGCAGTG 1849
1956 TGCCGAGCACAACAGCAGACAGATTCAGAGGGGCGCACTACAGTGGAGCCCTTACACTCA 2015
1850 TTCTGAATTTGACAGCATCCCTTTCGTGGGAAATTTACAAGTGGAACATACCGGGG 1909
2016 AGTAGAAGTACGAGCTTATGCAAACTCTACTGTATGCGAGAGGATTTGATTTCTTCTT 2075
1910 AGGGGGCGTGAAGGCC---TGCTCGCTCAGTGCCTAGCGGAAGGCTTCACTTCTACAC 1966
2076 TTCTTTGTCAATTAAGTCAAGATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTTG 2135
1967 GGAGAGGGCGGACCGTGGTGACCGGGACACCTGCGCTCCAGACACGGTGGACATTTG 2026
2136 TATAGATGGGATATGTAGAGAGTTGGATGTGACAAATGCTTGGATCTGATGCTGTGTA 2195
2027 CGTCAGTGGCGAATCAACAGCTGGGGCTGGACCGAGTCTTGGGCTCCGACTCGGGGA 2086
2196 AGAGCTCTGTGGGTGTGTAACGGGAATTAATCAGCTGTCGACGATTCACAGGGGTCTTA 2255
2087 GACAAAGTCCGAGTGTGGGGGTGACGGGCAAGTGCCTGCGAGACCAATCGAGGGGCTTT 2146
2256 CACCAAGCACCAACCAACCAACAGTATTAACATGTGTCAACATTCCTCTGAGAGCCG 2315
2147 CAGCCGAGCCTCACTGGGGCGGGTACGAGGATGTGCTGTGGATTTCCAAAGGCTCCGT 2206
2316 GAGTATCCGCACTATGAATGAAGCTCTTACCTCTTACATTTCTGTGCGCAATGCCCT 2375
2207 CCACATCTTATCCAGGATCTGAACCTCTCTCTCAGTCACTTGGCCCTGGAAGGAGACCA 2266
2376 CAGAAGGTACTACCTGAATGGGCACTGGACCGTGGAGTGGCCCGGCGGTACAAATTTTC 2435
2267 GGAGTCCCTGCTGCTGGAGGGGCTGCCCGGAGACCCCGGAGCCCGTCTGCTCTAGC 2326

2436 GGGCACTACTTTTCGACTACAGAGCGGTCTATATATGAGCCCGAGAACTTAAATCGCTACTGG 2495
2327 TGGGACCACTTTTAACTACGAGACGGGGCCAGACCCAGGTCCAGAGGCTCGAAGCCCTGGG 2386
2496 ACCAACCAACGAGACACATGATTTGTGGAGCTGTCTGTTTTCAGGGAAGGAACCCGGGTGTTGC 2555
2387 ACCGATTAAATGATCTCTCATCTGTCATGTTGTGCCCC---GGACGAGCTGCTGCCCC 2443
2556 CTGGGAATACTCCATGCTGCTGCTGGGGGACCGAGAAAGAGCCCCCTGCGCCAGCCAGCTA 2615
2444 CCGTACCCCTTCAATATGCCCCCATCGCCGTGACTCGCTGCCCCCTACTCTCTGGACATA 2503
2616 CACTTTGGGCGCATCTGTGCGCTCTGAGTGTCTCGGTGCTCGGAGGGGACAGATGACCGT 2675
2504 TGGCGCCCTGGACCAAGTGTCTGG-----CCAGTGTGCAAGCGGTAGCCAGGTGACGGC 2557
2676 GAGAGAGGGCTGTACAGAGACCTGAAAGTTTCAAGTAAATATATGTCCTTCTGCAATCCCCA 2735
2558 GGTGAGTGTGCGCAACCAAGCTGGAAGCTCGCGCGGTGCGCCCCCACTACTGAGTGGCCA 2617
2736 GACACGACCTGTACGCGGGCTGGTGCCTTTCGCAAGTATCTGCTGTGCTCCAGAGCTGGTC 2795
2618 CAGCAAGCTGCCCAAAAGGCGAGCGGCTGCAACACGAGAGCTTTGCCCCCTCAGACTGGGT 2677
2796 CGTGGGGAATGTGAGTGTCTGAGTTCGAGTTCGAGTGTGGCGGGGTGCCAGAGCCGCTGGTCC 2855
2678 TGTAGGGAATGTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2737
2856 GCAGTGTGACACGGCGGGT---GCATATGACTCGAGGCGAGTCCCGGCGAGCTGTGCCCC 2912
2738 CGTGTGCGAGCGCGCGTCTGTGCGGCGGAGGAGGCGCTGGAACGACAGCGCATGCC 2797
2913 TCAGCTGTCTCTCCAGCAGGAGGCTGCAACTCTCAGAGCTGCCCACTGCAATGAG 2972
2798 GCAGCGCGCGCCACTGTACTGAGGCGCTGCAACGCGGCCCTGCGAGTGGG 2857
2973 CGCGGGCCCTGGGAGAGTGTCTACACACTGTGGGAAGGGGTGAGGAAGCGGGCAGT 3032
2858 GGCCCTGAGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2917
3033 GGCTGTGAAGACCAACACCCCTCGGCGAGAGCGAGTGTGCTGCCGAGGCTGTCTGTCAC 3092
2918 CCTTTGACAGAGCGAGACCAACCGGCCAGCTGCCCCCGGCGCACTGTCTACCCGCGC 2977
3093 C 3093
2978 C 2978

RESULT 10

US-09-981-953A-1
; Sequence 1, Application US/09981953A
; Patent No. 6685599
; GENERAL INFORMATION:
; APPLICANT: RACIE, LISA A.
; APPLICANT: TWINE, NATALIE C.
; APPLICANT: AGOSTINO, MICHAEL J.
; APPLICANT: WOLFMAN, NEIL
; APPLICANT: MORRIS, ELISABETH A.
; TITLE OF INVENTION: NOVEL AGGRECANASE MOLECULES
; FILE REFERENCE: 08702.0075-00000
; CURRENT APPLICATION NUMBER: US/09/981,953A
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/242,317
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3766
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Nucleotide

QY 2913 TCAGCTGCTCCCTCCAGCAGGCGGCTCGAATCTTCAGAGCTGCCACCTCGATGGAG 2972
Db |||||
QY 3041 GCAGCCGCGCCACCTCTACTGAGGCGCTGCCAGCGCCCACTTGGCCCTCGGAGTGGG 3100
Db |||||
QY 2973 CGCCGGCCCTTGGCAGAGTGCTCACACACTGTGGGAAGGGGTGGAGGAAGCGGCAGT 3032
Db |||||
QY 3101 GGCCTCGACTGCTGAGTGCAACCCAGCTGCGGGCCGGGCTCCGCCACCAGCGTGGT 3160
Db |||||
QY 3033 GGCCTGTAAGACCAACCCCTCGGCCAGAGCGCAGCTGCTGCCGAGCGCTCTCTGCAC 3092
Db |||||
QY 3161 CCTTTGCAAGAGCGCAGACCAACCGCGCCAGCTGCCCGCGCACTGCTCACCCGCCG 3220
QY 3093 C 3093
Db 3221 C 3221

RESULT 11

US-09-963-791-23
; Sequence 23, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the Sa
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; PRIOR FILING DATE: 2000-12-08
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 2274
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-963-791-23

Query Match 7.2%; Score 266.2; DB 4; Length 2274;
Best Local Similarity 50.1%; Pred. No. 1.2e-69;
Matches 871; Conservative 0; Mismatches 803; Indels 63; Gaps 6;

QY 865 GAAGTGAACGTGGAGACCTTGGTGGTGGTCGACAAAGAGATGATGCAAAACCATGGCCAT 924
Db |||||
QY 292 GAACGGTTTGTGGAGACATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 351
Db |||||
QY 925 GAAATATATCACCACTTACGTCTCAGGATACCTCAACATGGTATCTGCTTTTATCAAGAT 984
Db |||||
QY 352 AAGACATTTGAACATTTAGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 411
Db |||||
QY 985 GGAACAATAGGAGAAACATCAACATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 1044
Db |||||
QY 412 TCCAGCTAGGAAACGTTGTGAATATTATAGTGGCCCGCTTAATTTGTTCTCACAGAAGAT 471
Db |||||
QY 1045 CAGCAGGAGTGGTATAGTCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1104
Db |||||
QY 472 CAGCCAACTTGGAGATAAACCACTATGAGATGATGATGATGATGATGATGATGATGATGATG 531
Db |||||
QY 1105 CAGTCTGGATTGAT-----GGGGAAGATGGGACT 1134
Db |||||
QY 532 CAGAAATCCATTCTCTCCACCAAGATGATGGAACACCATTTCCAGAAATGGGATGCC 591
Db |||||
QY 1135 CGTGATGACCAAGCAGTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1194
Db |||||
QY 592 CACCAAGATAATGACAGTTCTTTATTAAGTATGATGATGATGATGATGATGATGATGATGATG 651
Db |||||
QY 1195 TGTGACACTTTGGGATTTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1254
Db |||||

Db 652 TGTGGAACACTGGGCTTGGCCTCTCTGTGGCTGGAATGTGTGAGCCTGAAAGGAGCTGCAGC 711
QY 1255 ATTAATGAAGATACAGGCTCTTGGACTTGGCTTTCACCATTTGCCATGAGTCTCGACACACAC 1314
Db |||||
QY 712 ATTAATGAAGATCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 771
Db |||||
QY 1315 TTTGCGATGATTCATGATGAGGAAGGAACATGTGTAAAGTCCGAG-----GGC 1365
Db |||||
QY 772 TTTGATGAAACCATGATGGAATTTGAAATTTCTTGGGACGAAAGGTCATGAAGCAGCA 831
Db |||||
QY 1366 AACATCATGTCCCTACATTTGGCAGAGCAATGAGAGTCTTCTCTGTGTCACCTGCGAGC 1425
Db |||||
QY 832 AAACATTAATGAGCAGCTCACATTAATGCGAATACCAATTCCTTTTCTGCTGCTGCTGCTGCTG 891
Db |||||
QY 1426 CGCCAGTATCTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCTTCTGCTGATCAGCA 1485
Db |||||
QY 892 CGAGATACATCACCAGCTTTCTAGATTCAGCGCTGGTACTTTCCTGCTGATTAAGAGCC- 950
Db |||||
QY 1486 AAGCCTGTGAAGGAATACAAATGATCTCTGAGAAATTTGCCAGAGAAATTTATGATCAAAAC 1545
Db |||||
QY 951 --TCCCAAGCGTGAATTTCTTTTATCCAGCTGTGGCCCGAGGTCAAGTGTATGATGCTGAT 1008
Db |||||
QY 1546 ACACAGTCAAGTGGCAGTTTCGGAGAGAAAGCCAGCTCTGATGCTGCTGCTGCTGCTGCTG 1605
Db |||||
QY 1009 GAGCAATGTCTTTCCAGTATGGAGCAACCTCCCGCCCAATGTAATATG-----G 1059
Db |||||
QY 1606 GACATCTGTAAGCCCTGTGTGCTCATCTTATTTGGAAGAAATTTGAGACTTAAATTTATG 1665
Db |||||
QY 1060 GAAGTGTGTAGAGAGCTCTGTGTCTCAGCAAAAGCAACCGCTGTGTGTCACCAACAGTATT 1119
Db |||||
QY 1666 CCAGCAGCAGAGCAGCAATTTTGGGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1725
Db |||||
QY 1120 CCAGCAGCTGAGGGGACACTGTGTCAAACTGGGAATATTGAAAGAGGGTGTGTTATCAG 1179
Db |||||
QY 1726 AAATATGCTGATGAAG-----GCCCAAGCCCAACCCATGCGCACTGTGCGAGCTGG 1776
Db |||||
QY 1180 GGAGATTTGTTCTTTTGGCACTTGGCCCGAGAGCATAGATGGGGCTGGGCTCCCTGG 1239
Db |||||
QY 1777 TCTTCTGCTCCCATGCTCCAGGACCTGCGAGAGGGGAGTATCTCATAGAGTGCCTC 1836
Db |||||
QY 1240 TCACTATGGGAGAGTGCAGCAGGACCTGCGGGGAGGCGTCTCTCATCCCTAAGACAC 1299
Db |||||
QY 1837 TGCACCAACCCCAAGCCATCGCATGGAGGGAAGTTCTGTGAGGGCTCCATCGCAGCTCTG 1896
Db |||||
QY 1300 TGTGACAGTCCAGCACCTTCAGGAGGTGGAATAATTTGCTTTGGGGAAGAAACGGTAT 1359
Db |||||
QY 1897 AAGCTCTGCAACAGTCAAGAAATGTCCCGGACAGTGTGTGACTTCCGCTGCTGCTGCTGCTG 1956
Db |||||
QY 1360 CGCTCTGTAACACAGATCCATGCCCTTTGGGTTCCCGAGATTTTCGAGAGAAACAGTGT 1419
Db |||||
QY 1957 GCCGAGCACACAGCAGACGATTCAGAGGGGGGACCTACAGTGGGAAGCCCTTACACTCAA 2016
Db |||||
QY 1420 GCAGACTTTGACATAATGSCCTTTCCGAGGAAGATTAATTAATCGAAGAACCTTACTACT- 1476
Db |||||
QY 2017 GTAGAAGATCAGGACTTATGCAAACTCTACTGTATCGCAGAGGATTTGATTTCTTCTTT 2076
Db |||||
QY 1477 GGAGTGGGTAAACCTTGTGCAATTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1536
Db |||||
QY 2077 TCTTTGTCAAATAAAGTCAAGATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTGT 2136
Db |||||
QY 1537 GAACGTCTCTCGGCTGATCGATGGGACCCAGTCAATGCGGATTCACCTGGATATCTGC 1596
Db |||||
QY 2137 ATAGATCGGATATGTGAGAGAGTTGGATGTGCAATGCTCTTGGATCTGATGCTGTTGAA 2196
Db |||||
QY 1597 ATCAATGAGAAATGCAAGCAGCTAGGCTGTGATAATTTTGGGATCTGATGCTAGGGA 1656
Db |||||
QY 2197 GAGCTCTGTGGGTGTGTAAACGGGAATAACTCAGCCTGCAAGATTCACAGGGGTCTCTAC 2256
Db |||||
QY 1657 GATAGATGTGAGTCTGTGGAGGGGACGGAAGCAGATGTGATGCCATTTGAAGGTTCTTC 1716
Db |||||
QY 2257 ACCAAGCACACACACCAACAGTATATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2316
Db |||||
QY 1717 AATGATTTCACTGCCCCAGGGAGGCTTACATGGAAGTGTGTCAGATACCAAGAGGCTCTGTT 1776
Db |||||

QY 2317 AGTATCCGATCTATGAATGAACGCTCTTACCTCCTCAATTTCTGTGCGCAATGCCCTC 2376
Db |||||
2230 CACATGAAGTTAGAGAAATGGCCATGTCAAGAACTATATTTGCTTTAAAAATCTGAAGGA 2289
QY 2377 AGAAGGTACTACCTGAATGGGCACCTGGACCTGTGACCTGGCCCGCGGTACAAATTTTCG 2436
Db |||||
2290 GATGATTACTATATTAATGGTGCTGGACTATGACTGGCCCTAGGAAATTTGATTTGCT 2349
QY 2437 GGCACCTTTTCGACTACAGACGCTCTTATTAATGAGCCGAGAACTTTAAATCGCTACTGGA 2496
Db |||||
2350 GGGACAGCTTTTTCATTACAGAGACCAACTGATGAACAGAAATCTTTGGGAAGCTTAGGT 2409
QY 2497 CCAACCAACGAGACACTGATTTGGAGCTGCTTTTCAGGAGGAACCCGGGTGT 2553
Db |||||
2410 CCTACCTCAGAAATCTCATGTCATGTTCTGCTTCAAGAACAGAAATTTGGGAAT 2466

RESULT 13

US-09-369-364A-6
; Sequence 6, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Aptec, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 3218
; TYPE: DNA
; ORGANISM: Homo sapiens ADAMTS-7
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)..(3003)
US-09-369-364A-6

Query Match 6.9%; Score 255; DB 3; Length 3218;
Best Local Similarity 51.9%; Pred. No. 3.8e-66;
Matches 869; Conservative 0; Mismatches 755; Indels 51; Gaps 11;

QY 877 GAGACCTTGGTGGTGTCGACAAAAGATGATGCAAAACATCGGCCATGAAATATCACC 936
Db |||||
745 GAGACCTTGGTGGTGTCGACAAAAGATGATGCAAAACATCGGCCATGAAATATCACC 804
QY 937 ACCTAGCTGTCACGATCTCAACATGGTATCTGCTTTATTCAAAGATGGAAACATAGGA 996
Db |||||
805 AGCTATGTGTGACCATCATGAAATGATGGTGGTGGCTGCTTTTCATGACCCAGCATTTGGG 864
QY 997 GGAAACATCAACATTCGAATTTAGTGTCTGATTTCTTCTAGAGATGAAACCCAGGACTG 1056
Db |||||
865 AACCCTCATCCATCATCACCATTGTGGCTGTGCTGCTGGAAGATGAGAGGAGGACCTA 924
QY 1057 GTGATAAGTTCACGACGACACACCTTAAAGTAGCTTCTGCGAGTGGCAGCTGGAATTG 1116
Db |||||
925 AAGATCAGCACCATCATGACAAACACCTTGAAAGATGATGGAGCTTCTGCAAGTGGCAGAAAGCATC 984
QY 1117 ATGGGGAAGATGGGACTCGT-----CATGACACGCACTTACTGACTGGTCTG 1167
Db |||||
985 AACATGAAGGGGATGGCCATCCCTGCAACCATGACACTGGCCATCTCTCTCACCAGAAAG 1044
QY 1168 GATATATGTTCTGGAAGATGAGCCCTGTGACATCTTTGGGATTTGCAACCCATTAAGTGA 1227
Db |||||
1045 GACCTGTGTGACGACCATGAACCGGCCCTGTGAGACCTTGGGACTGTCCCATGTGGCGGGC 1104
QY 1228 ATGTGTAGTAAATATCGAGCTGACGATTAATGAGATACAGTCTTTGGACTGGCCTTC 1287
Db |||||
1105 ATGTCCAGCCGACCGCAGCTGACGATCAACGAGGACACGGGCTGCGCCTGCGCTTC 1164

QY 1288 ACCATTGCCATGAGTCTCGACACAACTTTGGCATGATTCATGATGGAGAGGGAACATG 1347
Db |||||
1165 ACTGTAGCCACAGCTCGGGCACAGTTTGGCAATTCAGCATGACGGAAGCGCAATGAC 1224
QY 1348 TGTAAAAAGTCCGAG-----GGCAACATCATGTCCCCTACATTTGGCAGGACGCAATGGA 1401
Db |||||
1225 TGTGAGCCCGTTGGGAAACGACCTTTTCATCATGTCTCCACAGCTCTCTGTACGACCGCT 1284
QY 1402 GTCTTCTCTCTGCTCACTCTGAGCCGCGCATGTATCTACAAATTTCTTAAGCAACCGCTCAA 1461
Db |||||
1285 CCGCTCAGCTGTCTGCTGAGCCGCGCATGTATCACAGGTTCTCTTGAACCGTGGGG 1344
QY 1462 GCTATCTGCTTGTGATCAGCCAAAGCCTGTGAAGGAATAACAAGTATCTCTGAGAAATG 1521
Db |||||
1345 GGCCTGTGCTGAGACGACCTCTG-----CCAAGGACATATTCGACTTCCCCTCGGTGCA 1401
QY 1522 CCAGAGAAATATATGATGCAAAACA CAGTGTCAAGTGGCAGTTCCGAGAGAAAGCCAAAG 1581
Db |||||
1402 CCGTGGCTCTCTATGATGTAAAGCCACAGTGTCCGCTCCAGTACGGGCTCTACTCTGCC 1461
QY 1582 CTCTGATGCTGGACTTTTAAAAAGGACATCTGTAAAGCCCTGTGTGTCATCGTATTGGA 1641
Db |||||
1462 TTCTGC-----GAGGACATGATAATGTCTGCCACACACTCTGTGTCTCTG---TGGGG 1512
QY 1642 AGGAATGTGAGACTTAATTTATGACAGCAGAGGACAAATTTGTGGGCATGACATG 1701
Db |||||
1513 ACCACTGTCTCTCAAGCTGGATGAGCTGTGACGGCACCCCGTGTGGGGAGAAATAG 1572
QY 1702 TGGTCCGGGAGGACAGTGTGAAATATGATGATGAGCCGCCCAAGCCCAACCCATGGC 1761
Db |||||
1573 TGGTGTCTCACTGGGAGTGGCT---ACCCGTGGGGCTTCGGGGCCCGAGGGCGTGGATGGT 1629
QY 1762 CACTGGTCCGACTGCTTCTTGTGTCCTCATGCTCCAGGACCTCGCGAGGGGAGTATCT 1821
Db |||||
1630 GGTGGTCTGGCTGAGCGCTGTCTCATCTGCTCACGAGCTGTGGCATGGGGGTACAG 1689
QY 1822 CATAGGAGTGGCTCTGACCAACCCCAAGCATCGCATGAGAGGGAAGTTCTGTGAGGGC 1881
Db |||||
1690 AGCCCGAGCGGAGTGCACGAGCTTACGCCCCAAATACAAAGGACAGATATGTGTGGGT 1749
QY 1882 TCCACTCGCACTCTGAAGCTCTGCAACAGTCAGAAATGTCCCGGGACAGTGTGTGACTTC 1941
Db |||||
1750 GAGCGAAGCGCTTCCGCTCTGCAACTGAGGCTGCTGCTGTGCGCCCTCTCTTC 1809
QY 1942 CGTCTGCTCAGTGTGCGGAGCACAACAGCAGACGATTCAGAGGGGCGCATCAAGTGG 2001
Db |||||
1810 CGCCAGCTCCAGTGCAGCCACTTTGACGCTATGCTCTTACAAGGGCCAGCTGCACACATGG 1869
QY 2002 AAGCCTTTACACTCAAGTAGAAGATCAGGACTTATGCAAACTCTACTGTATCGCAGAGGA 2061
Db |||||
1870 GTGCCCGTG-----GTCAATGAGTGAAACCTTGCAGCTGCACCTGCCGCGCGGAAAT 1923
QY 2062 TTTGATTTCTTTCTTTTCTTTTCAATAAAGTCAAAAGATGGGACTCCATGCTCGGAG--- 2118
Db |||||
1924 GAGTACTTTGCCAAGAGCTGCGGAGCGCTGTGTGATGCTGAGTGGCACCCTCTCTACCAGTTC 1983
QY 2119 ---GATAGCGCTAATTTCTATAGATGGATGTGTGAGAGAGTTGGATGTGACATGTC 2175
Db |||||
1984 CGAGCCAGCCGGGACCTCTGCATCAACGGCATCTGTAAGAAACGTGGGCTGTGATTCGAG 2043
QY 2176 CTTGGATCTGATCTGTGAAGACGCTGTGGGGTGTGTAACCGGGAATAACTCAGGCTGC 2235
Db |||||
2044 ATTGACTCCGGTCTATGGAGGACCGCTGTGGTGTGTGCCACCGCAACGGCTCCACTGC 2103
QY 2236 ACCATTCAAGGGTCTCTACACCAAGCACCAACCAACAGTATTATCAATGGTTC 2295
Db |||||
2104 CACACCGTGAAGCGGACCTTCG---AGGAGCGCGAGGGTCTGGGGTATGTGGATGTGGGG 2160
QY 2296 ACCATTCTCTGAGCCCGGAGATTCGCGCATCTATGAATGAACCTCTCTACTCTCTAC 2355
Db |||||
2161 CTGATCCAGCGCGGCGCACGCGAGATCCGATCCAGAGAGTTGCCGAGGCTGCCAACTTC 2220
QY 2356 ATTTCTGTGGCA---ATGCCCTCAGAAGGTACTACCTGAATGGGCACTGGACCGTGGAC 2412

Db 2221 CTGGCACTGGGAGCGAGGACCCGAGAAAGTACTTCTCAATGGTGGCTGGACCAATCCAG 2280
Qy 2413 TGGCCCGCGCGGTACAAATTTTCGGGCACCTACTTTTCGACTACAGACGGTCTTATAATGAG 2472
Db 2281 TGGAAACGGGACTACAGGTTGGCAGGACCACTTTCACATACCGACGCGGGGCAACTGG 2340
Qy 2473 CCGCAGAACTTAATCGGTACTGACCAACCAACGAGACACTGATGTGAGTGTC 2527
Db 2341 ---GAGAACTCACTGCTCCCGGGTCCCAACCAAGGAGCTGTCTGGATCCAGGTGC 2392

RESULT 14

US-09-369-364A-16
; Sequence 16, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 16
; LENGTH: 3885
; TYPE: DNA
; ORGANISM: Homo sapiens ADAMTS-10
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3246)
; NAME/KEY: misc_feature
; LOCATION: (3877)
; OTHER INFORMATION: n=A
; NAME/KEY: misc_feature
; LOCATION: (3882)
; OTHER INFORMATION: n=C
US-09-369-364A-16

Query Match 6.8%; Score 248.8; DB 3; Length 3885;
Best Local Similarity 48.7%; Pred. No. 3.4e-64;
Matches 1087; Conservative 0; Mismatches 1067; Indels 78; Gaps 12;
Qy 864 AGAACTGAACGTGGAGACCTTGGTGTGTCGACAAAAGATGATCAAAACCATGGCCA 923
Db 645 AGAGCGCTACGTGGAGACCATGATGTGCTGACAGATGATGGTGGCTTATCACGGCG 704
Qy 924 TGAATAATCACCACTACGTGCTACGATACCTCAACATGGTATCTGCTTTATTCAAAGA 983
Db 705 CCGGGATGTGGACGACTATGTCTCGGCCATCATGAACATTTGTGCCAAACTTTTCCAGGA 764
Qy 984 TGGAAATAGGAGGAAACATCAATTCGAATTTAGTCTGATTTCTTAGAAGATCA 1043
Db 765 CTGAGTCTGGGAAGACCGTTTAAACATCTCGTAACTCGCCCTCATCTGCTCACGGAGGA 824
Qy 1044 ACAGCCAGGACTGGTGATTAAGTCAACCAACGACACCACTTAAGTAGCTTCTGCGAGTG 1103
Db 825 CCAGCCCACTCTGGAGATCACCAACCATGCGGGAGTCCCTAGACAGCTTCTGTAAGTG 884
Qy 1104 GCAG-----TCTGGATTGAATGGGAAAGATGGGAC 1133
Db 885 GCAGAAATCCATCGTGAAACACAGCGGCCATGCAATGCCATTCACAGAGAACGGTGTGC 944
Qy 1134 TGTGATGACACGCCCATCTTACTGATCGTGTGGATATATGTTCTCGGAAGATGAGCC 1193
Db 945 TAAACCATGACACAGCAGTGTCTATCAACGCTATGACATCTGCACTACAGAAACAAACC 1004
Qy 1194 CTGTGACACTTTGGGATTTGACCCCAATAGTGAATGTGTAGTAAATATCGCAGCTGCAC 1253
Db 1005 CTCGGGCACACTAGGCTTGGC-CCGGTGGCGGAATGTGTGAGCGCGGAGAGAGCTGCAG 1063

Qy 1254 GATTAAATGAAGATACAGGTCTTGGACTGGCCTTTCACATTTGCCCATGAGTCTGGACACAA 1313
Db 1064 CGTCAATGGAGCAATTTGG-CTGCCACAAGCGTTTACCAATTTG-CCACGAGATCGGGACAC 1121
Qy 1314 CTTTGGCATGATTCATGATGGAGGAGGAAACATGTTGTAATAA-----AGTCCGAGGG 1364
Db 1122 ATTTCGCATGAACCATGACCGGTGGGAAACAGCTGTGGGGCCCGTGGTTCAGGACCCAGC 1181
Qy 1365 CAACATCATGTCTCCCTPACATTTGGCAGGAGCGAATGAGAGTCTTCTCTCGTGGTCACTTCGAG 1424
Db 1182 CAGCTCATGGCTGCCCAATTTACCATGAAGACCAACCCCATTCGTGTGGTTCATCTCTGCAA 1241
Qy 1425 CCGCCAGTATCTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCTTGTCTGATCAGCC 1484
Db 1242 CCGTCACTACATCACCAAGCTTTCTAGACTCGGGCTCTGGGGCTCTGCTGTAACCAACCGGCC 1301
Qy 1485 AAGCCTGTGAAGGAATAACAGTATCTCTGAGAAATTTGCCAGGAGAAATATATCATGCAAA 1544
Db 1302 ---CCCAGACAGGACTTTGTGTACCAGACAGTGGCACCGGGCCAGCCCTACGATGAGA 1358
Qy 1545 CACACAGTCAAAGTGGCAGTTCGGAGAGAAAGCCAAAGCTCTGCAATGCTGAGCTTTAAAAA 1604
Db 1359 TGAGCAATCGCGCTTTTCAGCATGGAGTCAAAATCGCGTCAAGTGA-----AATACGG 1409
Qy 1605 GGACATCTGTAAGCCCTGTGTGCCATCTGATTGGAAGAAATGTGAGACTAAATTTAT 1664
Db 1410 GGAGGTCTGACGCGACTGTGTGTCTGAGCAAGCAACCGGTGCAATCAACCAACAGCAT 1469
Qy 1665 GCCAGCAGCAGAGGACAAATTTGTGG-----CATGACATGTGTGCGGGGAGG 1715
Db 1470 CCGCGCGCGAGGGCAACGCTGTGCCAGACGCAACCATCGCAAGGGTGGTGTACAA 1529
Qy 1716 ACAGTGTGTGAATATGATGAAGGCCCAAGCCCAAGCCCATCGGCTGTCGAGCTG 1775
Db 1530 ACGGGTCTGTGTCCTTTGGGTGCGGCCAGAGGGTGTGACGAGGCTTGGGGGCGGTG 1589
Qy 1776 GTCTTCTTGGTCCCATGCTCAGGACCTGCGGAGGGGAGTATCTCAATGAGTGCCT 1835
Db 1590 GACTCATGGGCGACTGCGAGCGGACCTGTGCGGCGGCTGTCTCTTCTAGTGTCA 1649
Qy 1836 CTGCACCAACCCCAAGCCATCGATGAGGGAGTCTGTGAGGGTCTCACTCGCACTCT 1895
Db 1650 CTGCGACAGCCCCAGGCCAAACCATCGGGGGCAAGTACTGTCTGGGTGAGAGAGCGGCA 1709
Qy 1896 GAAGCTCTGCAACAGTCAAGAAATGTCCCGGACAGTGTGACTTCGCTGCTCAGTG 1955
Db 1710 CGCTCTCTGCAACACGAGTACTGTCTCCCTGGCTCCAGAGCTTCAGAGAGTGCAGTG 1769
Qy 1956 TCCGAGACACACAGCAGACGATTCAGAGGGCGGCACTACAGTGAAGCCTTTACACTCA 2015
Db 1770 TGTGAAATTTGACAGCATCTCTTTCGTGGGAAATTTCTACAAAGTGGAAACGTTACGGGG 1829
Qy 2016 AGTAGAAGTACAGGCTTATGCAAACTCTACTGATCGCAGAGGATTTGATTTCTTCTT 2075
Db 1830 AGGGGGCGTGAAGGGC---TGCTGCTCACAGGCTTAGCGGAAGGCTTCACTTCTACAC 1886
Qy 2076 TTCTTTGTCAAAATAAGTCAAAAGATGGGACTCCATGCTCGGAGGATAGCCGTAATCTTTG 2135
Db 1887 GGAGAGGGGGCGGACCGCTGGTGGACGGGACACCTCTGCGCTCCAGACACCGGTGGCAATTG 1946
Qy 2136 TATAGATGGATATGTGAGAGATGGATGTGACAAATGCTTGGATCTTGGATCTGATGTTGA 2195
Db 1947 CGTCAAGTGGGAAATGCAAGCAGTGTGGGCTGCGACCGAGTCTCTGGGCTCCGACCTGCGGGA 2006
Qy 2196 AGACGTCTGTGGGGTGTGTAAACGGGAATACTCAGCTGCAAGTTCACAGGGGTCTCTA 2255
Db 2007 GGCAAGTGTCCGAGTGTGTGGCGGTGAACGGAGTGTCTGCGAGACCATCGAGGGCGCTCTT 2066
Qy 2256 CACCAAGCACCAACACACCAACAGTATTTATCAATGGTCACTTCTTCTTGGAGCCCG 2315
Db 2067 GAGCCAGCTCACCTGGGGCGGGTACGAGGATGTCTGATTTCCCAAGGCTCCGT 2126
Qy 2316 GAGTATCGCATCTATGAAATGAACGCTCTTCTTACCTCTTACATTTTCTGTGCGCAATGCCCT 2375

```

Db 2127 CCACATCTTCATCCAGGATCGAACCTCTCTCTCAGTCACCTTGCCCTCGAAGGGAGACCA 2186
QY 2376 CAGAAGGTACTACCTGAATGGGCACTGGACCGTGGACTGGCCCGCGGTACAAATTTTC 2435
Db 2187 GGAGTCCCTGCTGCTGGAGGGGCTGCCTGGGACCCCCCAGCCCTCGCTCTAGC 2246
QY 2436 GGGCACTACTTTTGCACACTACAGACGGTCTTATATGAGCCCGAGAACTTATCGCTACTGG 2495
Db 2247 TGGGACCACTTTTCAACTGCGACAGGGGGCCAGACAGGTCCAGAGCCTCGAAGCCCTGGG 2306
QY 2496 ACCAACCAACGAGACACTGATTTGGGAGCTGCTGTTTCAGGAGAGGAACCCGGGTGTTC 2555
Db 2307 ACCGATTAATGCACTCTCTCATCGTCACTGCTGCTGGCCC---GGACCGAGCTGCTGCCCC 2363
QY 2556 CTGGGAATATCTCCATGCTGCTTGGGGACCGAAGCAGCCCTGTCGCCAGCCCACTA 2615
Db 2364 CCGCTACCGCTTCAATGCCCACTCGCCCGTGACTGCTGCCCCCTACTCTCTGGCACTA 2423
QY 2616 CACTTGGGCGCATCGTGGCTCTGAGTCTCGGTGCTCTGCGAGGGGGAAGATACCGGT 2675
Db 2424 TGCCTCTGGACCAAGTGTCTGG-----CCCACTGTGCGAGCGGTAGCCAGGTGCGAGGC 2477
QY 2676 GAGAGAGGCTGTACAGAGACCTGAAAGTTTCAAGTAAATATGCTCTCTGCAATCCCA 2735
Db 2478 GGTGGAGTGGCGCAACCCAGCTGGACAGCTCGCGGTGCGCCCCCACTACTGCGAGTCCCA 2537
QY 2736 GACAGGACCTGTACCGGGGCTGTGCTTGTCAAGATATCTGCTGCTCTCCAGCTGTGTC 2795
Db 2538 CAGCAAGCTGCCCAAGGCGGCGCTGCAACACGAGCCCTTGCCCTCCAGACTGGGT 2597
QY 2796 CGTGGGAACTGGAGTGCCTGCACTGCGAGTGGAGTGTGGGGGTGCCAGAGCGCCCGT 2855
Db 2598 TGTAGGAACTGTGCTCTGCGAGCGAGCTGCGATGCGAGCGTGGCGAGTACGTGCGT 2657
QY 2856 GCACTGCAACGGCGGT---GCACTATGACTCGGAGCCAGTCCCGGCCAGCTGTGCC 2912
Db 2658 CGTGTGCGAGCGCGGCTCTCTGCGCGGAGGAGAGCGCTGGACGACAGCGCATGCC 2717
QY 2913 TCAGCTGCTGCTCCTCAGCAGGCGAGCGCTGCAACTCTCAGAGCTGCCACCTGCAATGAG 2972
Db 2718 GCAGCGCGGCCACTGTACTGGAGCGCTGCCACGCGCCCACTTGCCCTCCGAGTGGGC 2777
QY 2973 CGCGGGCCCTGGCGAGAGTGTCAACACCTGTGGGAAAGGGGTGGAGGAAGCGGCGAGT 3032
Db 2778 AACCTCTGAGTGTCTGAGTGTACCCCAAGCTGTGGCGCTGCTCCGSCCAACGAGTGT 2837
QY 3033 GGCCTGTGAAG 3044
Db 2838 CCTTGTGAAG 2849

```

```

RESULT 15
US-09-369-364A-12
; Sequence 12, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 5804
; TYPE: DNA
; ORGANISM: Homo sapiens ADAMTS-9
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(5648)

```

```

; NAME/KEY: misc feature
; LOCATION: (1406)
; OTHER INFORMATION: n=T
; NAME/KEY: misc feature
; LOCATION: (1563)
; OTHER INFORMATION: n=T
US-09-369-364A-12

```

```

Query Match      4.4%; Score 160.4; DB 3; Length 5804;
Best Local Similarity 48.9%; Pred. No. 3.7e-37;
Matches 682; Conservative 0; Mismatches 673; Indels 39; Gaps 8;

QY 867 ACTGAACCTGGAGACCTTGGTGGTGGACAAAAGATGACAAAACCCAGTGGCCATGA 926
Db 722 ACGGTTTGTAGAGCTTGGTGGTGGCAGACAAACAGATGGTTTCATACCATGG---AGA 778
QY 927 AAATATCACCACTTACAGTCTCAGATACTCAACATGGTATCTGCTTTATTCAAAGATGG 986
Db 779 AAACCTTCAACACTATATTTTAACTTTAATGTCAATTTAGCTCTATCTATATAAGACCC 838
QY 987 AACAAATAGAGGAAACATCAACATTTGCAATTTGAGTCTGATTTCTTAGAAGATGAACA 1046
Db 839 AAGTATTTGGAATTTAATTAATTTATTTGTTGTAACCTTAATTTGATTCATAATGAACA 898
QY 1047 GCCAGGACTGGTATAGTCAACGACAGACCAACCTTTAAGTAGTCTTCCAGTGGCA 1106
Db 899 GGATGGGCTTCCATATCTTTTAAATGCTCAGACAACTTAAACACTTTTGGCAGTGGCA 958
QY 1107 GTCTGGATTTGATGGGAAAGATGGGACTGCTGATGACCAACCGCATCTTACTGACTGGTCT 1166
Db 959 GCAAT---TCGAACAGTCCAGGTGGAATCCATCATGATATCTGCTTTCTTTAAACAAGACA 1015
QY 1167 GGATATATGTTTCTCGAAGAAATGAGCCCTGTGACACTTTTGGGATTTGCACCCATAAGTGG 1226
Db 1016 GGATATCTG---CAGAGCTCAGCAAAATGTGATACCTTAGGCTGSGCTGACTGGGAAC 1072
QY 1227 AATGTGTAGTAAATATCGCAGCTGACAGATTAATGAAGATACAGGTCTTGAGCTGGCCTT 1286
Db 1073 CATTTGTGATCCCTTATAGAAGCTGTTCTATTTAGTGAAGATAGTGGATTTGATGACAGCTTT 1132
QY 1287 CACATTTGCCCATGAGTCTGGACAACTTTTGGCATGATTCATGATGCGAGAGGGAACAT 1346
Db 1133 TACGATCGCCCATGAGCTGGGCCATGTGTTTAACTGCTCTCATGATGACAAACAATG 1192
QY 1347 GTGTAAAGTCCGAGGGCA-----ACATCATGTCCCTTACATTTGGCAGGACGCAA 1397
Db 1193 TAAAGAAAGAGGAGTTAAGAGTCCCGAGCATGTCTGCTCCAACACTGAACTTCTACAC 1252
QY 1398 TGGAGTCTTCTCTGGTCAACCCTGCGAGCGCCAGTATCTACAAATTTTAAAGCACC 1457
Db 1253 CAACCCCTGGATGTGGTCAAGTGTAGTCAAAATATATCATCTGAGTTTTTTAGACACTGG 1312
QY 1458 TCAAGCTATCTGCTTGTCTGATCAGCCAAAGCCTGTGAAGGAATACAAGTATCTCTGAGAA 1517
Db 1313 TTATGGCGAGTGTGCTTAAACGAACTGATCC---AGACCCCTACCCCTTGGCTGTCCA 1369
QY 1518 ATTGCCAGGAGAATTATATGATGCAACACACAGTGCAGTGGCAGTTCGGAGAGAAAGC 1577
Db 1370 ACTGCCAGGCACTCTTTTACAACTGATTAACAAATNGAATTTGATTTTTTGGACAGGTTTC 1429
QY 1578 CAAGCTCTGCACTGCTGGACTTTTAAAGAGACATCTCTTAAAGCCCTCTGTTGGCCATCGTAT 1637
Db 1430 TCAGGTGTGCCCATATATGATGTCAGTGCAGACGGCTCTGGTGCATTAACGTCAATGAGT 1489
QY 1638 TGAAGGAATGTGAGACTTAATTTTATGCCAGCAGAGGAGGACAAATTTTGGGGCATGA 1697
Db 1490 ACACAAAGGCTGCCGAGCTCAGCACACACCTGGGCGCGTGGGAGCGAGTGGAGCGCTGG 1549
QY 1698 CATGTGTGCGGGGAGGACAGTGTCTGTAATAATATGTTGATGAGGCCCCCAAGCCCA 1757
Db 1550 AAGCACTGCAGNATGGATTTTGTGTTCCCAAGAAATGGATGTCCCCGTG---ACAGA 1606
QY 1758 TGCCCACTGGTGGACTGGTCTTCTTTGGTCCCATGCTCCAGGACCTCGGAGGGGAGT 1817

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2005, 20:57:22 ; Search time 4963 Seconds
(without alignments)
4590.278 Million cell updates/sec

Title: US-10-804-457-3
Perfect score: 3675
Sequence: 1 atgaagcccgccgcccgcgg.....gtcttaagtccaacttgtga 3675

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3675	100.0	3675	13	US-10-217-774-3
2	3675	100.0	3675	21	US-10-804-457-3
3	3675	100.0	4042	13	US-10-217-774-5
4	3675	100.0	4042	21	US-10-804-457-5
5	3673.4	100.0	3675	15	US-10-296-616-1
6	3671.8	99.9	3675	17	US-10-330-176-1
7	3671.8	99.9	3675	18	US-10-275-107-16

8	3135.2	85.3	4888	17	US-10-399-645-19	Sequence 19, Appli
9	2055.6	55.9	2997	10	US-09-981-151A-1	Sequence 1, Appli
10	2022.8	55.0	2902	10	US-09-981-151A-5	Sequence 5, Appli
11	1874.6	51.0	2895	10	US-09-981-151A-7	Sequence 7, Appli
12	1449.8	39.5	1476	13	US-10-217-774-1	Sequence 1, Appli
13	1449.8	39.5	1476	21	US-10-804-457-1	Sequence 1, Appli
14	1139.4	31.0	3666	16	US-10-240-545A-1	Sequence 1, Appli
15	1135.6	30.9	3663	18	US-10-354-983-1	Sequence 1, Appli
16	1031.8	28.1	2805	18	US-10-275-107-18	Sequence 18, Appli
17	1019.2	27.7	3219	18	US-10-354-983-3	Sequence 3, Appli
18	1010.6	27.5	3389	18	US-10-363-937-34	Sequence 34, Appli
19	988.2	26.9	3246	19	US-10-391-364-88	Sequence 88, Appli
20	988.2	26.9	3445	19	US-10-391-364-86	Sequence 86, Appli
21	978.2	26.6	2433	10	US-09-981-151A-3	Sequence 3, Appli
22	900.8	24.5	3630	17	US-10-161-493-123	Sequence 123, App
23	861.6	23.4	2589	14	US-10-226-560-1	Sequence 1, Appli
24	861.6	23.4	3013	14	US-10-226-560-3	Sequence 3, Appli
25	860	23.4	2940	17	US-10-188-186-131	Sequence 131, App
26	835	22.7	3555	17	US-10-188-186-129	Sequence 129, App
27	666.8	18.1	3954	18	US-10-354-983-5	Sequence 5, Appli
28	622	16.9	1986	18	US-10-354-983-7	Sequence 7, Appli
29	558.8	15.2	1059	17	US-10-161-493-129	Sequence 129, App
30	557.2	15.2	1059	17	US-10-161-493-131	Sequence 131, App
31	557.2	15.2	1059	17	US-10-161-493-133	Sequence 133, App
32	555.6	15.1	1059	17	US-10-161-493-125	Sequence 125, App
33	555.6	15.1	1059	17	US-10-161-493-127	Sequence 127, App
34	421.4	11.5	2217	14	US-10-226-560-6	Sequence 6, Appli
35	421.2	11.5	476	10	US-09-918-995-15335	Sequence 15335, A
36	405.6	11.0	1530	14	US-10-226-560-4	Sequence 4, Appli
37	359.2	9.8	624	17	US-10-188-186-155	Sequence 155, App
38	323.8	8.8	525	16	US-10-029-386-8689	Sequence 8689, Ap
39	322.8	8.8	348	16	US-10-029-386-22417	Sequence 22417, A
40	317	8.6	5338	9	US-09-788-043C-6	Sequence 6, Appli
41	316.6	8.6	5070	20	US-10-164-893-2	Sequence 2, Appli
42	316.6	8.6	5115	17	US-10-295-027-1100	Sequence 1100, Ap
43	316.6	8.6	5115	20	US-10-783-528-31	Sequence 31, Appli
44	316.6	8.6	5115	21	US-10-651-237-15	Sequence 15, Appli
45	316.6	8.6	5115	21	US-10-782-413-15	Sequence 15, Appli

ALIGNMENTS

RESULT 1
US-10-217-774-3
; Sequence 3, Application US/10217774
; Publication No. US20020193583A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. US20020193583A1el Human Proteases and Polynucleotides Encodi
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/10/217,774
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US/09/930,872
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3675
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-217-774-3

Query Match 100.0%; Score 3675; DB 13; Length 3675;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3675; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ATGAAGCCCGCCGCGCGGCTTGGCGGCGCTGTGGATGCTGTGGCGCAG 60
|||||

Db 1 ATGAAGCCCGCGCGCGGATGGCGGGCTTGGCGGCGCTGTGTGGATCTGCTGGCGCAG 60
Qy 61 GTGGCCGAGCAGGCACTTGTGCGTGCCTATGGGACCCGAGCGGCGAGCGCCTGGGAGCCG 120
Db 61 GTGGCCGAGCAGGCACTTGTGCGTGCCTATGGGACCCGAGCGGCGAGCGCCTGGGAGCCG 120
Qy 121 AGCGTCCGCGCTCTCTCCACCCGCGGAGCGCGCGGCTGTGATGGAAAGGGCGAATAT 180
Db 121 AGCGTCCGCGCTCTCTCCACCCGCGGAGCGCGCGGCTGTGATGGAAAGGGCGAATAT 180
Qy 181 GACCTGGTCTCTGCTACGAGTTGACCAAGGGGCGATTAACGTGTCCCATGAATCATG 240
Db 181 GACCTGGTCTCTGCTACGAGTTGACCAAGGGGCGATTAACGTGTCCCATGAATCATG 240
Qy 241 CACCATCAGCGCGGAGAGAGCAGTGGCCGTGTCCGAGTTGAGTCTCTTCCACCTTCGG 300
Db 241 CACCATCAGCGCGGAGAGAGCAGTGGCCGTGTCCGAGTTGAGTCTCTTCCACCTTCGG 300
Qy 301 CTGAAAGGCTCCAGGCAGCACTTCCAGTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360
Db 301 CTGAAAGGCTCCAGGCAGCACTTCCAGTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360
Qy 361 CCTGGCTTATGTGTCAGAGCTTGGGAAAGCAGGCACTAAGTCTGTGCGAGCTTTACCG 420
Db 361 CCTGGCTTATGTGTCAGAGCTTGGGAAAGCAGGCACTAAGTCTGTGCGAGCTTTACCG 420
Qy 421 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGCGATCACAGAAACTCCTCAGTGGCC 480
Db 421 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGCGATCACAGAAACTCCTCAGTGGCC 480
Qy 481 CTTTCAACCTGCCAAGGCTTGTGAGCATGATACGAACAGAGAGGAGATTAATTCCTTA 540
Db 481 CTTTCAACCTGCCAAGGCTTGTGAGCATGATACGAACAGAGAGGAGATTAATTCCTTA 540
Qy 541 AGGCCACTTCTTCCACACTCTCATGTGAACCTCGGAGAGCTGCCAAGGCGAGCTCGCCA 600
Db 541 AGGCCACTTCTTCCACACTCTCATGTGAACCTCGGAGAGCTGCCAAGGCGAGCTCGCCA 600
Qy 601 TCCGAGTACTGTACAGAGATCCACAGAGCCCATGCTCTGGGCGCAGTGAAGTCTCTG 660
Db 601 TCCGAGTACTGTACAGAGATCCACAGAGCCCATGCTCTGGGCGCAGTGAAGTCTCTG 660
Qy 661 GTGACCTCAAGGACATGGGAGCTGGCAGCATCAACCCCTGCGACAGAGGACCTTGGCCTG 720
Db 661 GTGACCTCAAGGACATGGGAGCTGGCAGCATCAACCCCTGCGACAGAGGACCTTGGCCTG 720
Qy 721 GGAATGCCCAAAAGCAGCATTTCTGTGAAGACGCAAGAAATACATGCCCGAGCTCCC 780
Db 721 GGAATGCCCAAAAGCAGCATTTCTGTGAAGACGCAAGAAATACATGCCCGAGCTCCC 780
Qy 781 AAGGAAGACCTCTTTCATCTTGCAGATGAGTATAGTCTTGTACGGCATAAAGCGTCT 840
Db 781 AAGGAAGACCTCTTTCATCTTGCAGATGAGTATAGTCTTGTACGGCATAAAGCGTCT 840
Qy 841 CTTCTGAGTCCCATAGAAATGAAGAACTGAACGTTGGAGACTTGTGTGGTGGTGGTGGT 900
Db 841 CTTCTGAGTCCCATAGAAATGAAGAACTGAACGTTGGAGACTTGTGTGGTGGTGGTGGT 900
Qy 901 AAGATGATGCAAAACCAATGGCCATGAAGAAATATACCACTACGCTGCTCAGCATCAAC 960
Db 901 AAGATGATGCAAAACCAATGGCCATGAAGAAATATACCACTACGCTGCTCAGCATCAAC 960
Qy 961 ATGGTATCTGCTTTATTCAAGATGGAACATAGAGGAAACATCAACATTTGCAATTTGTA 1020
Db 961 ATGGTATCTGCTTTATTCAAGATGGAACATAGAGGAAACATCAACATTTGCAATTTGTA 1020
Qy 1021 GGTCTGATCTTCTAGAGATGAACAGCAGGACTGTGATTAAGTCAACCGCAGACCC 1080
Db 1021 GGTCTGATCTTCTAGAGATGAACAGCAGGACTGTGATTAAGTCAACCGCAGACCC 1080
Qy 1081 ACCTTAAGTAGCTTCTGCGAGTGGCAGTCTGATTTGATGGGAAAGTGGACTCGTCAT 1140
Db 1081 ACCTTAAGTAGCTTCTGCGAGTGGCAGTCTGATTTGATGGGAAAGTGGACTCGTCAT 1140

Qy 1141 GACCACGCCATCTTACTGACTGGTCTGGATATATGTTCTTGAAGAAATGAGCCCTGTGAC 1200
Db 1141 GACCACGCCATCTTACTGACTGGTCTGGATATATGTTCTTGAAGAAATGAGCCCTGTGAC 1200
Qy 1201 ACTTTGGGATTTGCACCCATAAGTGGAAATGCTAGTAAATATATCGCAGCTGACAGTAAT 1260
Db 1201 ACTTTGGGATTTGCACCCATAAGTGGAAATGCTAGTAAATATATCGCAGCTGACAGTAAT 1260
Qy 1261 GAAGATACAGGCTTTGGACTGGCTTCCATTTGACCTGAGTCTGAGCAACAATTTGGC 1320
Db 1261 GAAGATACAGGCTTTGGACTGGCTTCCATTTGACCTGAGTCTGAGCAACAATTTGGC 1320
Qy 1321 ATGATTCATGATGGAGAAAGGAAACATGTGTAAAGTCCGAGGGCAACATCATGTCCCT 1380
Db 1321 ATGATTCATGATGGAGAAAGGAAACATGTGTAAAGTCCGAGGGCAACATCATGTCCCT 1380
Qy 1381 ACATTTGGCAGGACGCAATGGAGTCTTCTCTGGTCCACCTGAGCCGCGAGTATCTACAC 1440
Db 1381 ACATTTGGCAGGACGCAATGGAGTCTTCTCTGGTCCACCTGAGCCGCGAGTATCTACAC 1440
Qy 1441 AAATTTCTAAGCACCGCTCAAGCTATCTGCTTCTGATCAGCCAAAGCCTGTGAAGAA 1500
Db 1441 AAATTTCTAAGCACCGCTCAAGCTATCTGCTTCTGATCAGCCAAAGCCTGTGAAGAA 1500
Qy 1501 TACAGTATCTCTGAGAAATTTGCCAGGAGAAATATATGATGCAAAACACACAGTGAAGTG 1560
Db 1501 TACAGTATCTCTGAGAAATTTGCCAGGAGAAATATATGATGCAAAACACACAGTGAAGTG 1560
Qy 1561 CAGTTCCGAGGAAAGCCAGCTCTGATGCTGGAATTTTAAAGGACATCTGTAAAGCC 1620
Db 1561 CAGTTCCGAGGAAAGCCAGCTCTGATGCTGGAATTTTAAAGGACATCTGTAAAGCC 1620
Qy 1621 CTGTGTGCCATCGTATTTGGAAGAAATGTGAGACTAAATTTATGCCAGCAGCAGAGGC 1680
Db 1621 CTGTGTGCCATCGTATTTGGAAGAAATGTGAGACTAAATTTATGCCAGCAGCAGAGGC 1680
Qy 1681 ACATTTGTGGGATGATGATGCTGGTCCGGGAGGACAGTGTGTGAAATATGTTGATGAA 1740
Db 1681 ACATTTGTGGGATGATGATGCTGGTCCGGGAGGACAGTGTGTGAAATATGTTGATGAA 1740
Qy 1741 GGCCCCAAGCCCAACCCATGCGACTTGTGCGACTGTCTTCTTGGTCCCGAGCTCCAG 1800
Db 1741 GGCCCCAAGCCCAACCCATGCGACTTGTGCGACTGTCTTCTTGGTCCCGAGCTCCAG 1800
Qy 1801 ACCTGCGGAGGGGAGTATCTCATAGAGTGGCTCTGACCAACCCCAAGCCATCGCAT 1860
Db 1801 ACCTGCGGAGGGGAGTATCTCATAGAGTGGCTCTGACCAACCCCAAGCCATCGCAT 1860
Qy 1861 GGAGGGAAAGTTCTGTGAGGCTCCACTGCGACTGTGAGTCTTCTTGGTCCCGAGCTCCAG 1920
Db 1861 GGAGGGAAAGTTCTGTGAGGCTCCACTGCGACTGTGAGTCTTCTTGGTCCCGAGCTCCAG 1920
Qy 1921 CCCCGGACAGTGTGACTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Db 1921 CCCCGGACAGTGTGACTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Qy 1981 AGAGGGCGGACTTACAGGTGGAAGCTTACACTCAAGTAGAAGATCAGGACTTATGCAAA 2040
Db 1981 AGAGGGCGGACTTACAGGTGGAAGCTTACACTCAAGTAGAAGATCAGGACTTATGCAAA 2040
Qy 2041 CTCTACTGATCGCAGAGGATTTGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2100
Db 2041 CTCTACTGATCGCAGAGGATTTGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2100
Qy 2101 GGGACTCCATGCTCCGAGGATAGCCGTAATGTTGTATAGATGGGATATGTGAGAGATT 2160
Db 2101 GGGACTCCATGCTCCGAGGATAGCCGTAATGTTGTATAGATGGGATATGTGAGAGATT 2160
Qy 2161 GGATGTGCAATGTCTTGGATCTGATGCTGTGAAGACGCTGTGGGGTGTGTAAACGG 2220
Db 2161 GGATGTGCAATGTCTTGGATCTGATGCTGTGAAGACGCTGTGGGGTGTGTAAACGG 2220

```
QY 2221 AATAACTCAGCTGACGATTTCACAGGGTCTCTACACCAAGCACCAACCAACCAACCAAG 2280
Db 2221 AATAACTCAGCTGACGATTTCACAGGGTCTCTACACCAAGCACCAACCAACCAACCAAG 2280
QY 2281 TATTATCAGATGTCACCAATTCCTCTGAGCCCGGAGTATCGCATCTATGAAATGAAC 2340
Db 2281 TATTATCAGATGTCACCAATTCCTCTGAGCCCGGAGTATCGCATCTATGAAATGAAC 2340
QY 2341 GTCTCTACCTCTACATTTCTGCGCAATGCGCTCAGAAAGTACTTCTGAAATGGCAC 2400
Db 2341 GTCTCTACCTCTACATTTCTGCGCAATGCGCTCAGAAAGTACTTCTGAAATGGCAC 2400
QY 2401 TGGACCGTGGATGCGCCCGGCGGTACAAATTTTCGGGCACACTTTTCGACTACAGACG 2460
Db 2401 TGGACCGTGGATGCGCCCGGCGGTACAAATTTTCGGGCACACTTTTCGACTACAGACG 2460
QY 2461 TCCATAATAGAGCCCGAGAACTTAATCGCTACTGGAACCAACCAACGAGACACTGATTG 2520
Db 2461 TCCATAATAGAGCCCGAGAACTTAATCGCTACTGGAACCAACCAACGAGACACTGATTG 2520
QY 2521 GAGCTGCTTTTTCAGGGAAGAACCCGGGTGTTGCTGGGAATPACTCCATGCTCGCTTG 2580
Db 2521 GAGCTGCTTTTTCAGGGAAGAACCCGGGTGTTGCTGGGAATPACTCCATGCTCGCTTG 2580
QY 2581 GGGACCGAGAACGAGCCCTGCGCCAGCCAGCTACACTTGGGCCATCGTGGCTCTGAG 2640
Db 2581 GGGACCGAGAACGAGCCCTGCGCCAGCCAGCTACACTTGGGCCATCGTGGCTCTGAG 2640
QY 2641 TGCTCGCTGCTCTGCGAGGGGAGACAGATGACCGTGAGAGAGGGCTGTACAGAGACCTG 2700
Db 2641 TGCTCGCTGCTCTGCGAGGGGAGACAGATGACCGTGAGAGAGGGCTGTACAGAGACCTG 2700
QY 2701 AAGTTTCAAGTAATAATGCTTCTGCAATCCCAAGACGACCTGTGTACGGGGCTGGTG 2760
Db 2701 AAGTTTCAAGTAATAATGCTTCTGCAATCCCAAGACGACCTGTGTACGGGGCTGGTG 2760
QY 2761 CCTTGGAAAGTATCTGCTGCTCCAGCTGCTCCAGCTGCTGGGGAACTGGAGTGCCTGAGT 2820
Db 2761 CCTTGGAAAGTATCTGCTGCTCCAGCTGCTCCAGCTGCTGGGGAACTGGAGTGCCTGAGT 2820
QY 2821 CGGAGCTGTGGGGGGTCCAGAGCGCCCGCTGCTGAGTGCACACGGGGGTGCACTAT 2880
Db 2821 CGGAGCTGTGGGGGGTCCAGAGCGCCCGCTGCTGAGTGCACACGGGGGTGCACTAT 2880
QY 2881 GACTCGGAGCCAGTCCCGGCCAGCTGTGCTCCTCAGCTGCTCCCTCCAGAGCAGGCGC 2940
Db 2881 GACTCGGAGCCAGTCCCGGCCAGCTGTGCTCCTCAGCTGCTCCCTCCAGAGCAGGCGC 2940
QY 2941 TGGAACTCTCAGAGTGCACACCTGATGAGGGCGCGGCTGAGAGTGGGAGAGTGTCTACAC 3000
Db 2941 TGGAACTCTCAGAGTGCACACCTGATGAGGGCGCGGCTGAGAGTGGGAGAGTGTCTACAC 3000
QY 3001 ACCTGTGGGAAGGGTGGAGGAAGCGGGCAGTGGCTGTAAAGCAACCAACCCCTCGGC 3060
Db 3001 ACCTGTGGGAAGGGTGGAGGAAGCGGGCAGTGGCTGTAAAGCAACCAACCCCTCGGC 3060
QY 3061 AGAGCGCAGCTGTGCGCCAGCCTGTCTGACCTCCGAGCCCAAGCCCAAGAGTGCATGAA 3120
Db 3061 AGAGCGCAGCTGTGCGCCAGCCTGTCTGACCTCCGAGCCCAAGCCCAAGAGTGCATGAA 3120
QY 3121 GCTGTCTGCTTACGGCTGCGCAACGCGCTGCTGAGTGCAGTGCAGTGGTGTGCTCGCTGG 3180
Db 3121 GCTGTCTGCTTACGGCTGCGCAACGCGCTGCTGAGTGCAGTGCAGTGGTGTGCTCGCTGG 3180
QY 3181 TCCAGTGTCTGTGATGTGAAGAGGAACACAGAAAGATTTCTTAAATGTGCTGAA 3240
Db 3181 TCCAGTGTCTGTGATGTGAAGAGGAACACAGAAAGATTTCTTAAATGTGCTGAA 3240
QY 3241 AAGTATGTTTCTGGAAAGTATCGAGAGCTGGCTCAAGAAAGTGTCTCAATTTGCGAAG 3300
Db 3241 AAGTATGTTTCTGGAAAGTATCGAGAGCTGGCTCAAGAAAGTGTCTCAATTTGCGAAG 3300
QY 3301 CCCAGCTGAGCTGGAAAGCTGCTCGCGCCGCTTCCATGCGCCAGGCAACCCCAATTT 3360
```

```
Db 3301 CCACAGCTGAGCTGGAACGCTGCGCCCGCTTCCATGCCAGGCAACCCCAATTT 3360
QY 3361 GCTGCTCGGGACCCCTCGAGGGGCGAGCTGTTTGCCTCACCCCTGGTCTCAGTGCACGGCC 3420
Db 3361 GCTGCTCGGGACCCCTCGAGGGGCGAGCTGTTTGCCTCACCCCTGGTCTCAGTGCACGGCC 3420
QY 3421 AGCTGTGGGGAGGGCTTCAGACAGAGTCCGTGCGAGTCTGCTGGGGCGCGCCGGCC 3480
Db 3421 AGCTGTGGGGAGGGCTTCAGACAGAGTCCGTGCGAGTCTGCTGGGGCGCGCCGGCC 3480
QY 3481 TCAGGCTGCTCTGTCACACAGAGCTTCGGGCTCCCTGGCTGCAACACTCACTTCTGC 3540
Db 3481 TCAGGCTGCTCTGTCACACAGAGCTTCGGGCTCCCTGGCTGCAACACTCACTTCTGC 3540
QY 3541 CCCATTGACAGAGAAGAAAGATGCTTCTGCAAAAGACTACTTCCACTGGTGTACCTGGTA 3600
Db 3541 CCCATTGACAGAGAAGAAAGATGCTTCTGCAAAAGACTACTTCCACTGGTGTACCTGGTA 3600
QY 3601 CCCCAGCAGGGATGTGACAGCCACAAGTTCTACGGCAAGCAGTGTGCAAGACTTGTCT 3660
Db 3601 CCCCAGCAGGGATGTGACAGCCACAAGTTCTACGGCAAGCAGTGTGCAAGACTTGTCT 3660
QY 3661 AAGTCCAACTTGTGA 3675
Db 3661 AAGTCCAACTTGTGA 3675

RESULT 2
US-10-804-457-3
; Sequence 3, Application US/10804457
; Publication NO. US20050065334A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the
; FILE REFERENCE: Same
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/10/804,457
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: US/10/217,774
; PRIOR FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US/09/930,872
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3675
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-804-457-3

Query Match 100.0%; Score 3675; DB 21; Length 3675;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3675; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCGCGCGCGGATGGCGGGCTTGGCGGCGCTGTGGATGCTGTGGCGCAG 60
Db 1 ATGAGAGCCCGCGCGCGGATGGCGGGCTTGGCGGCGCTGTGGATGCTGTGGCGCAG 60
QY 61 GTGGCCGAGCAGGCACTTCGCTGCGCCATGGGACCCGAGCGGCGAGCGCTCGGAGCCCG 120
Db 61 GTGGCCGAGCAGGCACTTCGCTGCGCCATGGGACCCGAGCGGCGAGCGCTCGGAGCCCG 120
QY 121 AGCTGCCCGCTCTCTTCCACCCCGGAGCGCGCGCTGTGATGGAAGGCGGAATAT 180
Db 121 AGCTGCCCGCTCTCTTCCACCCCGGAGCGCGCGCTGTGATGGAAGGCGGAATAT 180
QY 181 GACCTGTGCTCTGCTGTGAGTTGACCAAGGGGCGATTACGTGCTCCCATGAATCATG 240
Db 181 GACCTGTGCTCTGCTGTGAGTTGACCAAGGGGCGATTACGTGCTCCCATGAATCATG 240
```


Db 2401 TGAACGTGACCTGGCCCGCCGGTCAAAATTTTCGGGCACTACTTTTCGACTACAGCGG 2460
QY 2461 TCCTATAATAGAGCCCGAGAACTTAATCGCTACTGGACCAACCAACGAGACACTGATTGTG 2520
Db 2461 TCCTATAATAGAGCCCGAGAACTTAATCGCTACTGGACCAACCAACGAGACACTGATTGTG 2520
QY 2521 GAGCTGCTGTTTCAGGGAAGAAACCGGGTGTGTCCTGGGAATCTCCATGCTCGCTTG 2580
Db 2521 GAGCTGCTGTTTCAGGGAAGAAACCGGGTGTGTCCTGGGAATCTCCATGCTCGCTTG 2580
QY 2581 GGGACCGAGAGCAGCCCTGCCCCAGCCAGCTACACTTTGGGCCATCGTGCCCTCTGAG 2640
Db 2581 GGGACCGAGAGCAGCCCTGCCCCAGCCAGCTACACTTTGGGCCATCGTGCCCTCTGAG 2640
QY 2641 TGCTCCGTGTCTCGGAGGGGAGCAGATGACCGTGGAGAGGGGCTGCTACAGAGACCTG 2700
Db 2641 TGCTCCGTGTCTCGGAGGGGAGCAGATGACCGTGGAGAGGGGCTGCTACAGAGACCTG 2700
QY 2701 AGTTTCAAGTAATAATATGCTCTCTGCAATCCCAAGACACGACTGTGTCAGGGGCTGGTG 2760
Db 2701 AGTTTCAAGTAATAATATGCTCTCTGCAATCCCAAGACACGACTGTGTCAGGGGCTGGTG 2760
QY 2761 CCTTGCAAGTATCTGCTGTCTCCAGCTGCTCGTGGGAACTGGAGTGCCTGCAGT 2820
Db 2761 CCTTGCAAGTATCTGCTGTCTCCAGCTGCTCGTGGGAACTGGAGTGCCTGCAGT 2820
QY 2821 CGGACGTGTGGCGGGGTGCCAGAGCCGCCCGTGCAGTGCAACGGGGGTGCACTAT 2880
Db 2821 CGGACGTGTGGCGGGGTGCCAGAGCCGCCCGTGCAGTGCAACGGGGGTGCACTAT 2880
QY 2881 GACTCGGAGCACTCCCGGCAGCTGTGCTCCTCAGCTGCTGCTCCCTCCAGCAGGAGGCC 2940
Db 2881 GACTCGGAGCACTCCCGGCAGCTGTGCTCCTCAGCTGCTGCTCCCTCCAGCAGGAGGCC 2940
QY 2941 TGCAACTCTCAGAGCTGCCACCTGCAATGGAGCGCCGGGCTGGGAGAGTGTCTCACAC 3000
Db 2941 TGCAACTCTCAGAGCTGCCACCTGCAATGGAGCGCCGGGCTGGGAGAGTGTCTCACAC 3000
QY 3001 ACCTGTGGGAAGGGGTGGAGGAAGCGGCAAGTGGCTGTAAAGACCAACCCCTCGGCC 3060
Db 3001 ACCTGTGGGAAGGGGTGGAGGAAGCGGCAAGTGGCTGTAAAGACCAACCCCTCGGCC 3060
QY 3061 AGAGCGCAGCTGTGCTCCGACGCTGTGCACTCCGAGCCCAAGCCAGGATGCAAGAA 3120
Db 3061 AGAGCGCAGCTGTGCTCCGACGCTGTGCACTCCGAGCCCAAGCCAGGATGCAAGAA 3120
QY 3121 GCCTGTCTGCTTTCAGCGCTGCCAAGCCCAAGAGCTGCAGTGGCTGGTCCGCTGG 3180
Db 3121 GCCTGTCTGCTTTCAGCGCTGCCAAGCCCAAGAGCTGCAGTGGCTGGTCCGCTGG 3180
QY 3181 TCCACGTGCTCTGTGACATGTGAAGAGGAACACAGAAAGATTTCTTAAATGTGTGAA 3240
Db 3181 TCCACGTGCTCTGTGACATGTGAAGAGGAACACAGAAAGATTTCTTAAATGTGTGAA 3240
QY 3241 AAGTATGTTTCTGGAAAGTATCGAGAGCTGGCTCAAGAGAGTGTCTCAATTTGCCGAAG 3300
Db 3241 AAGTATGTTTCTGGAAAGTATCGAGAGCTGGCTCAAGAGAGTGTCTCAATTTGCCGAAG 3300
QY 3301 CCCAGCTGAGCTGGAACGTGCTGCGCCCGCTTCCATGCCCCAGGACCCCCCATTTT 3360
Db 3301 CCCAGCTGAGCTGGAACGTGCTGCGCCCGCTTCCATGCCCCAGGACCCCCCATTTT 3360
QY 3361 GCTGCTCGGGGACCTTCGAGGGGACGTGGTTTGGCTCAACCTGCTCAGTGCAGGGCC 3420
Db 3361 GCTGCTCGGGGACCTTCGAGGGGACGTGGTTTGGCTCAACCTGCTCAGTGCAGGGCC 3420
QY 3421 AGCTGTGGGGAGCGCTTCAGACAGAGTCCGTCAGTGCCTGCTGGCGGGGCCCGGCC 3480
Db 3421 AGCTGTGGGGAGCGCTTCAGACAGAGTCCGTCAGTGCCTGCTGGCGGGGCCCGGCC 3480
QY 3481 TCAGGCTGCTCTGCAACAGAAAGCTTTCGGCTCCCTGGCTGCAACACTCACTTCTGCTC 3540

Db 3481 TCAGGCTGCTCTGCAACAGAAAGCCTTCGGCTCCTGGCTGCAACACTCACTTCTGCTC 3540
QY 3541 CCCATTGCAGAGAGAAAGATGCTTTCTGCAAAAGACTACTTCCACTGGTCTACCTGGTA 3600
Db 3541 CCCATTGCAGAGAGAAAGATGCTTTCTGCAAAAGACTACTTCCACTGGTCTACCTGGTA 3600
QY 3601 CCCAGCAGGGATGTCAGGCCAACAGTTCTACGGCAAGCAGTGTCTGCAAGACTTCTCT 3660
Db 3601 CCCAGCAGGGATGTCAGGCCAACAGTTCTACGGCAAGCAGTGTCTGCAAGACTTCTCT 3660
QY 3661 AAGTCCAACTTTGTGA 3675
Db 3661 AAGTCCAACTTTGTGA 3675

RESULT 3
US-10-217-774-5
; Sequence 5, Application US/1021774
; Publication No. US20020193583A1
; GENERAL INFORMATION:
; APPLICANT: Fridde, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. US20020193583A1 Human Proteases and Polynucleotides Encodi
; FILE REFERENCE: Same
; CURRENT APPLICATION NUMBER: US/10/217,774
; PRIOR FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US/09/930,872
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4042
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-217-774-5

Query Match 100.0%; Score 3675; DB 13; Length 4042;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3675; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGCCCGCGCGCGATGGCGGGCTTGGCGCGCTGTGGATGCTGTGGCGCAG 60
Db 99 ATGAAGCCCGCGCGCGATGGCGGGCTTGGCGCGCTGTGGATGCTGTGGCGCAG 158
QY 61 GTGGCGGAGCAGGCACTGCTGCCATGGGACCCGCGCGCGCGCGCGCGCGCGCG 120
Db 159 GTGGCGGAGCAGGCACTGCTGCCATGGGACCCGCGCGCGCGCGCGCGCGCGCG 218
QY 121 AGCGTCCCGCTCTCTCCACCCGCGGAGCGCGCGCTGGATGGAAGGCGGAATAT 180
Db 219 AGCGTCCCGCTCTCTCCACCCGCGGAGCGCGCGCTGGATGGAAGGCGGAATAT 278
QY 181 GACCTGGTCTCTGCTACGAGGTTGACCAAGGGCGGATTAAGTGTCCCATGAATCATG 240
Db 279 GACCTGGTCTCTGCTACGAGGTTGACCAAGGGCGGATTAAGTGTCCCATGAATCATG 338
QY 241 CACCATCAGCGCGGAGAGAGCAGTGGCGGCTGTCGAGGTTGAGTCTTACCTTCGG 300
Db 339 CACCATCAGCGCGGAGAGAGCAGTGGCGGCTGTCGAGGTTGAGTCTTACCTTCGG 398
QY 301 CTGAAGGCTCCAGGCAACGCTTCCACGTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360
Db 399 CTGAAGGCTCCAGGCAACGCTTCCACGTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 458
QY 361 CTTGGCTTTATGTGTCAGACGTTGGGAAAGACAGGCACTAAGTCTGTGAGACTTTACCG 420
Db 459 CTTGGCTTTATGTGTCAGACGTTGGGAAAGACAGGCACTAAGTCTGTGAGACTTTACCG 518
QY 421 CAGAGGACTTCTGTTTCTATCAAGGCTTTTGGGATCAACAGAACTCTCAGTGGCC 480

Db	519	CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGCGATCACACAGAAACTCCTCAGTGGCC	578
Qy	481	CTTTCAACTGCGCAAGGCTTGTGAGCATGATACGAACAGAGAGGACAGATTACTTCTTA	540
Db	579	CTTTCAACTGCGCAAGGCTTGTGAGCATGATACGAACAGAGAGGACAGATTACTTCTTA	638
Qy	541	AGGCCACTTCTTTTCAACCTCTCATGAGAACTCGGCAGAGCTGCCAAGGCAGCTCGCCA	600
Db	639	AGGCCACTTCTTTTCAACCTCTCATGAGAACTCGGCAGAGCTGCCAAGGCAGCTCGCCA	698
Qy	601	TCCCAAGTACTGTACAAGAGATCCACAGAGCCCCATGCTCTGGGGCCAGTGAGTCTGTG	660
Db	699	TCCCAAGTACTGTACAAGAGATCCACAGAGCCCCATGCTCTGGGGCCAGTGAGTCTGTG	758
Qy	661	GTGACCTCAAGGCATGGGAGCTGGCACATCAACCCCTGCGACAGCGAGACTTCGCGCTG	720
Db	759	GTGACCTCAAGGCATGGGAGCTGGCACATCAACCCCTGCGACAGCGAGACTTCGCGCTG	818
Qy	721	GGACTGCCACAAAAGCAGCATTTCTGTGGAGCGCAGAATAATACATGCCCCAGGCTGCC	780
Db	819	GGACTGCCACAAAAGCAGCATTTCTGTGGAGCGCAGAATAATACATGCCCCAGGCTGCC	878
Qy	781	AAGGAAGACTCTTTCATCTTGCCAGATGAGTATAAGTCTTGCTTACGGCATAAGCGTCT	840
Db	879	AAGGAAGACTCTTTCATCTTGCCAGATGAGTATAAGTCTTGCTTACGGCATAAGCGTCT	938
Qy	841	CTTCTGAGGTCCCATAGAAATGAGAACTGAACTGGAGACTTGGTGGTGGTGGTGGTGGT	900
Db	939	CTTCTGAGGTCCCATAGAAATGAGAACTGAACTGGAGACTTGGTGGTGGTGGTGGTGGT	998
Qy	901	AAGATGATGCAAAACCATTGGCCATGAAATAATCACCACTACGTGCTCAGCATACTCAAC	960
Db	999	AAGATGATGCAAAACCATTGGCCATGAAATAATCACCACTACGTGCTCAGCATACTCAAC	1058
Qy	961	ATGATATCTGCTTTTATTCAAAGATGGAAACAATAGGAGGAAACATCAAATTCGCAATTGTA	1020
Db	1059	ATGATATCTGCTTTTATTCAAAGATGGAAACAATAGGAGGAAACATCAAATTCGCAATTGTA	1118
Qy	1021	GGTCTGATTTCTTGTAGAAGATGAACAGCAGGACTGTGTATAGTCAACACAGGAGCCAC	1080
Db	1119	GGTCTGATTTCTTGTAGAAGATGAACAGCAGGACTGTGTATAGTCAACACAGGAGCCAC	1178
Qy	1081	ACCTTAAGTAGCTTCTGCCAGTGGCAGCTTGGATGTAGTGGGAAAGATGGGACTCGTCAT	1140
Db	1179	ACCTTAAGTAGCTTCTGCCAGTGGCAGCTTGGATGTAGTGGGAAAGATGGGACTCGTCAT	1238
Qy	1141	GACACGCCATCTTACTGACTGTGCTGGATATATGTTCTGGAGATGAGCCCTGTGAC	1200
Db	1239	GACACGCCATCTTACTGACTGTGCTGGATATATGTTCTGGAGATGAGCCCTGTGAC	1298
Qy	1201	ACTTTGGGATTTGCAACCCATPAGTGGAAATGTAGTAAATATCGCAGCTCCACGATTAAT	1260
Db	1299	ACTTTGGGATTTGCAACCCATPAGTGGAAATGTAGTAAATATCGCAGCTCCACGATTAAT	1358
Qy	1261	GAAGATACAGGCTCTTGAGCTGGCTTCCACCATTTGCCCATGAGTCTGGACAATACTTTGGC	1320
Db	1359	GAAGATACAGGCTCTTGAGCTGGCTTCCACCATTTGCCCATGAGTCTGGACAATACTTTGGC	1418
Qy	1321	ATGATTCATGATGGAGAAAGGAAACATGTGTAAAAAGTCCGAGGGCAACATGTCGCCCT	1380
Db	1419	ATGATTCATGATGGAGAAAGGAAACATGTGTAAAAAGTCCGAGGGCAACATGTCGCCCT	1478
Qy	1381	ACATTTGGCAGGCGCAATGGAGTCTTCTCTGTGTCAACCTTGAGCGCCGCAATCTACAC	1440
Db	1479	ACATTTGGCAGGCGCAATGGAGTCTTCTCTGTGTCAACCTTGAGCGCCGCAATCTACAC	1538
Qy	1441	AAATTTCTAAGCAACCGCTCAAGCTATCTGCCTGTGTGATCGCCAAAGCTGTGAAGGAA	1500
Db	1539	AAATTTCTAAGCAACCGCTCAAGCTATCTGCCTGTGTGATCGCCAAAGCTGTGAAGGAA	1598
Qy	1501	TACAGTATCTTGAGAAATTCGACGAGAAATATATGATGCAAAACAACAGTGCAGTGG	1560
Db	1599	TACAGTATCTTGAGAAATTCGACGAGAAATATATGATGCAAAACAACAGTGCAGTGG	1658

```
QY 2641 TGCTCCGCTGCTCCTCGGAGGGGACACATGACCGTGAAGAGGGGCTGCTACAGACACCTG 2700
Db 2739 TGCTCCGCTGCTCCTCGGAGGGGACACATGACCGTGAAGAGGGGCTGCTACAGACACCTG 2798
QY 2701 AAGTTTCAAGTAAATATGCTCTCTGCAATCCCAAGACACAGCTGTCACGGGCTGGTG 2760
Db 2799 AAGTTTCAAGTAAATATGCTCTCTGCAATCCCAAGACACAGCTGTCACGGGCTGGTG 2858
QY 2761 CCTTGCAAAATATGCTCTCTCTGCAATCCCAAGACACAGCTGTCACGGGCTGGTG 2820
Db 2859 CCTTGCAAAATATGCTCTCTCTGCAATCCCAAGACACAGCTGTCACGGGCTGGTG 2918
QY 2821 CGGACGTGTGGGGGGTGCACAGAGCCGCGCTGCAAGTGCACAGCGCGGGTGCACAT 2880
Db 2919 CGGACGTGTGGGGGGTGCACAGAGCCGCGCTGCAAGTGCACAGCGCGGGTGCACAT 2978
QY 2881 GACTCGAGAGCAGTCCCGGCGAGCCTGTGCGCTCAGCTGCTCCCTCCAGCAGCAGGCC 2940
Db 2979 GACTCGAGAGCAGTCCCGGCGAGCCTGTGCGCTCAGCTGCTCCCTCCAGCAGCAGGCC 3038
QY 2941 TGAACCTCTCAGAGCTGCCCACTGTCATGAGCGCGCGCTGCGGCGAGTGTCTCACAC 3000
Db 3039 TGAACCTCTCAGAGCTGCCCACTGTCATGAGCGCGCGCTGCGGCGAGTGTCTCACAC 3098
QY 3001 ACTGTGGGAAGGGTGGAGAGCGGCGAGTGGCGCTGTGAAGACCAACCCCTCGGCC 3060
Db 3099 ACTGTGGGAAGGGTGGAGAGCGGCGAGTGGCGCTGTGAAGACCAACCCCTCGGCC 3158
QY 3061 AGAGCGCAGCTGTGCCCGCAGCTGTCTGACCTCCGAGCCCAAGCCAGCATGATGAA 3120
Db 3159 AGAGCGCAGCTGTGCCCGCAGCTGTCTGACCTCCGAGCCCAAGCCAGCATGATGAA 3218
QY 3121 GCCTGTCTGTGCTCAGCGCTGCCCAAGCCTGATGAGCGCGCGCTGCGGCGAGTGTCT 3180
Db 3219 GCCTGTCTGTGCTCAGCGCTGCCCAAGCCTGATGAGCGCGCGCTGCGGCGAGTGTCT 3278
QY 3181 TCCAGTGTCTGTGATGTGAAGAGGAAACACAGAAAGATCTTAAATGTGTGAA 3240
Db 3279 TCCAGTGTCTGTGATGTGAAGAGGAAACACAGAAAGATCTTAAATGTGTGAA 3338
QY 3241 AAGTATGTTTCTGGAAGTATCAGAGCTGGCGCTCAAGAGGTCACATTTGCGGAG 3300
Db 3339 AAGTATGTTTCTGGAAGTATCAGAGCTGGCGCTCAAGAGGTCACATTTGCGGAG 3398
QY 3301 CCCAGCTGTGAGTGTGAACGTGCTGCGCCCGCTTCATGCCCGCAGCACCCCGCATTT 3360
Db 3399 CCCAGCTGTGAGTGTGAACGTGCTGCGCCCGCTTCATGCCCGCAGCACCCCGCATTT 3458
QY 3361 GCTGCTCGGGAACCTTCGAGGGGAGCTGTTGTTTGCCTCACCTGGTCTCAGTGCAAGGCC 3420
Db 3459 GCTGCTCGGGAACCTTCGAGGGGAGCTGTTGTTTGCCTCACCTGGTCTCAGTGCAAGGCC 3518
QY 3421 AGCTGTGGGGGGGCGTTTCAAGCAGGTCCGTGAGTGGCTGAGCTGGGGCGCGCGGCC 3480
Db 3519 AGCTGTGGGGGGGCGTTTCAAGCAGGTCCGTGAGTGGCTGAGCTGGGGCGCGCGGCC 3578
QY 3481 TCAGGCTGCTCCTCGCACCAAGAGCCTTCGCGCTCCCTGCGCTGCAACACTCACTCTTGC 3540
Db 3579 TCAGGCTGCTCCTCGCACCAAGAGCCTTCGCGCTCCCTGCGCTGCAACACTCACTCTTGC 3638
QY 3541 CCCATTGAGAGAGAAAGATGCTTCTGCAAGAATCTTCACTTCACTGCTGCTACCTGGA 3600
Db 3639 CCCATTGAGAGAGAAAGATGCTTCTGCAAGAATCTTCACTTCACTGCTGCTACCTGGA 3698
QY 3601 CCCAGCAGGGATGTGACCAAGTCTACGGCAGCAGCTGCTGCAAGCTGCTGCTCT 3660
Db 3699 CCCAGCAGGGATGTGACCAAGTCTACGGCAGCAGCTGCTGCAAGCTGCTGCTCT 3758
QY 3661 AAGTCCAACTTGTGA 3675
Db 3759 AAGTCCAACTTGTGA 3773
```

RESULT 4

```
US-10-804-457-5
; Sequence 5, Application US/10804457
; Publication No. US20050065334A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/10/804,457
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: US/10/217,774
; PRIOR FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US/09/930,872
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4042
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-804-457-5
```

Query Match 100.0%; Score 3675; DB 21; Length 4042;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3675; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAAGCCCGCGCGCGGATGGCGGGCTTGGCGGCTGTGGATGCTGTGGCGCAG 60
Db 99 ATGAAGCCCGCGCGCGGATGGCGGGCTTGGCGGCTGTGGATGCTGTGGCGCAG 158
QY 61 GTGGCCGAGCAGCAGCTGCGCGCATGGGACCCGACGCGCAGCGCTGGAGGCCG 120
Db 159 GTGGCCGAGCAGCAGCTGCGCGCATGGGACCCGACGCGCAGCGCTGGAGGCCG 218
QY 121 AGCGTCCCGCGTCTCTCCACCCGCGAGCGCGCGGCTGGATGGAAGGCGCAATAT 180
Db 219 AGCGTCCCGCGTCTCTCTCCACCCGCGAGCGCGCGGCTGGATGGAAGGCGCAATAT 278
QY 181 GACCTGGTCTCTGCTTACGAGGTGACCAAGGGGCGATTAAGTGTGCCAATAATCATG 240
Db 279 GACCTGGTCTCTGCTTACGAGGTGACCAAGGGGCGATTAAGTGTGCCAATAATCATG 338
QY 241 CACCATCAGCGCGGAGAGCAGTGGCGGCTGTCGAGGTGAGTCTCTTCACTTCGG 300
Db 339 CACCATCAGCGCGGAGAGCAGTGGCGGCTGTCGAGGTGAGTCTCTTCACTTCGG 398
QY 301 CTGAAAGGCTCCAGGCGACGACTTCCACGTGGATCTCAGGACTTCCAGCAGCCTAGTGGCT 360
Db 399 CTGAAAGGCTCCAGGCGACGACTTCCACGTGGATCTCAGGACTTCCAGCAGCCTAGTGGCT 458
QY 361 CTTGGCTTTATTGTGAGACGTTTGGGAAAGACAGGCACTAAGTCTGTGACAGATTACCG 420
Db 459 CTTGGCTTTATTGTGAGACGTTTGGGAAAGACAGGCACTAAGTCTGTGACAGATTACCG 518
QY 421 CCAGGAGCTCTGTTTCTATCAAGGCTCTTTGGGATCACACAGAACTCTCAGTGGCC 480
Db 519 CCAGGAGCTCTGTTTCTATCAAGGCTCTTTGGGATCACACAGAACTCTCAGTGGCC 578
QY 481 CTTTCAACCTGCGCAAGGCTTGTCAAGCATGATACAGAGAGGCGAGATTACTTCTTA 540
Db 579 CTTTCAACCTGCGCAAGGCTTGTCAAGCATGATACAGAGAGGCGAGATTACTTCTTA 638
QY 541 AGGCCACTTCTTTCACACTCTCTCATGGAATCTGGCAGAGCTGCCCAAGGAGCTCGCCA 600
Db 639 AGGCCACTTCTTTCACACTCTCTCATGGAATCTGGCAGAGCTGCCCAAGGAGCTCGCCA 698
QY 601 TCCGAGTACTGTACAGAGATCCACAGAGCCCAATGCTCTCTGGGCGCAGTAGGTCTTG 660
Db 699 TCCGAGTACTGTACAGAGATCCACAGAGCCCAATGCTCTCTGGGCGCAGTAGGTCTTG 758
```

Qy	661	GTGACCTCAAGACATGGGAGCTGGCAATCAACCCCTGCACACGACGACGACCTTCGCCTG	720
Db	759	GTGACCTCAAGGACATGGGAGCTGGCAATCAACCCCTGCACAGCAGACCTTCGCCTG	818
Qy	721	GGACTGCCACAAAAGCAGCATTTCTGTGGAGAGCGAAGAAATACATGCCCCAGCCTCCC	780
Db	819	GGACTGCCACAAAAGCAGCATTTCTGTGGAGAGCGAAGAAATACATGCCCCAGCCTCCC	878
Qy	781	AGGGAAGACCTCTTCACTCTGCCAGATGAGTATAAGTCTCTGCTTACGGCATAGCGCTCT	840
Db	879	AGGGAAGACCTCTTCACTCTGCCAGATGAGTATAAGTCTCTGCTTACGGCATAGCGCTCT	938
Qy	841	CTTCTGAGGTCCTCATAGAAATGAAGACTGAAAGCTGGAGACCTTGGTGTGTGCACAAA	900
Db	939	CTTCTGAGGTCCTCATAGAAATGAAGACTGAAAGCTGGAGACCTTGGTGTGTGCACAAA	998
Qy	901	AAAGATGATGCAAAACCATGGCCATGAAATATCACCACTAGCTCTACGATACTCAAC	960
Db	999	AAAGATGATGCAAAACCATGGCCATGAAATATCACCACTAGCTCTACGATACTCAAC	1058
Qy	961	ATGGTATCTGCTTTATTCAAGAGTGAACAAATAGGAGGAAACATCAAATTGCAATTGTA	1020
Db	1059	ATGGTATCTGCTTTATTCAAGAGTGAACAAATAGGAGGAAACATCAAATTGCAATTGTA	1118
Qy	1021	GGTCTGATTTCTTAGAAGATGAAACAGCAGGACTGGTGAATGATCAACGCGAGACAC	1080
Db	1119	GGTCTGATTTCTTAGAAGATGAAACAGCAGGACTGGTGAATGATCAACGCGAGACAC	1178
Qy	1081	ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGATTTGATGGGGAAGATGGGATCTGCTAT	1140
Db	1179	ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGATTTGATGGGGAAGATGGGATCTGCTAT	1238
Qy	1141	GACACGCCATCTTACTGACTGGTCTGATATATGTTCTCGAAGAAATGAGCCCTGTGAC	1200
Db	1239	GACACGCCATCTTACTGACTGGTCTGATATATGTTCTCGAAGAAATGAGCCCTGTGAC	1298
Qy	1201	ACTTTGGGATTTGCCACCAATAGTGGAAATGTAGTAAATCCACGCTCCACGATTAAT	1260
Db	1299	ACTTTGGGATTTGCCACCAATAGTGGAAATGTAGTAAATCCACGCTCCACGATTAAT	1358
Qy	1261	GAAGATACAGGCTTTGGACTGGCCTTACCAATGCCCATGATGCTGGACACAACTTTGGC	1320
Db	1359	GAAGATACAGGCTTTGGACTGGCCTTACCAATGCCCATGATGCTGGACACAACTTTGGC	1418
Qy	1321	ATGATTTATGATGGGAAGAGGAAACATGTGTAAAAGTCCGAGGCGAACATCATGTCCCCT	1380
Db	1419	ATGATTTATGATGGGAAGAGGAAACATGTGTAAAAGTCCGAGGCGAACATCATGTCCCCT	1478
Qy	1381	ACATTTGGCAGGACGCAATGGAGTCTTCTCTGTGTCACCTGACGCGCCAGTATCTACAC	1440
Db	1479	ACATTTGGCAGGACGCAATGGAGTCTTCTCTGTGTCACCTGACGCGCCAGTATCTACAC	1538
Qy	1441	AAATTTCTAAGCAACGCTCAAGCTATCTGCTGCTGATCAGCCAAAGCTCTGGAAGAA	1500
Db	1539	AAATTTCTAAGCAACGCTCAAGCTATCTGCTGCTGATCAGCCAAAGCTCTGGAAGAA	1598
Qy	1501	TACAAGTATCTTGAAAAATTCGACAGAGAAATATATGATGCAAAACACAGTGCAGTGG	1560
Db	1599	TACAAGTATCTTGAAAAATTCGACAGAGAAATATATGATGCAAAACACAGTGCAGTGG	1658
Qy	1561	CAGTTCCGAGAGAAAGCCAGCTCTGCATCTGGAATTTAAAAAGGACATCTGTAAAGCC	1620
Db	1659	CAGTTCCGAGAGAAAGCCAGCTCTGCATCTGGAATTTAAAAAGGACATCTGTAAAGCC	1718
Qy	1621	CTGTGTGCCATCGTATTGGAAGAAATGTGAGACTTAAATTTATCCACGACGACGAGGC	1680
Db	1719	CTGTGTGCCATCGTATTGGAAGAAATGTGAGACTTAAATTTATCCACGACGACGAGGC	1778
Qy	1681	ACAATTTGTGGGCAATGCATGTGTGTCGGGGGAGGACGTGTGTAATAATGTGTGATGAA	1740
Db	1779	ACAATTTGTGGGCAATGCATGTGTGTCGGGGGAGGACGTGTGTAATAATGTGTGATGAA	1838

Qy	1741	GGCCCCAAGCCACCCATGSCCACTGGTCGGACTGGTCTTTCTTGGTCCCACATGCTCCAGG	1800
Db	1839	GGCCCCAAGCCACCCATGSCCACTGGTCGGACTGGTCTTTCTTGGTCCCACATGCTCCAGG	1898
Qy	1801	ACCTGCGGAGGGGGAGTATCTCATAGGAGTCCGCTCTGCACCAACCCCAAGCCATCGCAT	1860
Db	1899	ACCTGCGGAGGGGGAGTATCTCATAGGAGTCCGCTCTGCACCAACCCCAAGCCATCGCAT	1958
Qy	1861	GGAGGGAAGTTCCTGTGAGGGCTCCACTCGCACTCTGAGAGCTGTGCAACAGTCAAGAAATGT	1920
Db	1959	GGAGGGAAGTTCCTGTGAGGGCTCCACTCGCACTCTGAGAGCTGTGCAACAGTCAAGAAATGT	2018
Qy	1921	CCCCGGGACAGTGTGATCTTCGGTGTCTCAGTGTGCCGAGCACCAACAGCAGACGATTC	1980
Db	2019	CCCCGGGACAGTGTGATCTTCGGTGTCTCAGTGTGCCGAGCACCAACAGCAGACGATTC	2078
Qy	1981	AGAGGGCCGCGCACTACAAGTGGAGCCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAA	2040
Db	2079	AGAGGGCCGCGCACTACAAGTGGAGCCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAA	2138
Qy	2041	CTCTACTGTATCGCAGAGGATTTGATTTCTTCTTTCTTTGTCAAAATAAAGTCAAAAGAT	2100
Db	2139	CTCTACTGTATCGCAGAGGATTTGATTTCTTCTTTCTTTGTCAAAATAAAGTCAAAAGAT	2198
Qy	2101	GGGACTCCATGCTCGGAGGATAGCCGTAATGTTTGTATAGATGGGATATGTGACAGAGTT	2160
Db	2199	GGGACTCCATGCTCGGAGGATAGCCGTAATGTTTGTATAGATGGGATATGTGACAGAGTT	2258
Qy	2161	GGATGTGCAATGTCTTGGATCTGATGCTTTGAGAGCCGAGTATCCGACTCTGTGGGGTGTGAACGGG	2220
Db	2259	GGATGTGCAATGTCTTGGATCTGATGCTTTGAGAGCCGAGTATCCGACTCTGTGGGGTGTGAACGGG	2318
Qy	2221	AATAACTCAGCCTGCACGATTTCAAGGGGTCTCTACCAAGCACCAACCAACCCAG	2280
Db	2319	AATAACTCAGCCTGCACGATTTCAAGGGGTCTCTACCAAGCACCAACCAACCCAG	2378
Qy	2281	TATTATCAATGTCTACCAATTCCTCTGAGAGCCGAGTATCCGACTCTATGAAATGAAC	2340
Db	2379	TATTATCAATGTCTACCAATTCCTCTGAGAGCCGAGTATCCGACTCTATGAAATGAAC	2438
Qy	2341	GTCTCTACCTCTPACATTTCTGTGCCAATAGCCCTCAGAAAGTACTTACTGAAATGGGCAC	2400
Db	2439	GTCTCTACCTCTPACATTTCTGTGCCAATAGCCCTCAGAAAGTACTTACTGAAATGGGCAC	2498
Qy	2401	TGGACCGTGGACTGGCCGCGCGGTACAAATTTTGGGCACACTTTCAGCTACACAGCCG	2460
Db	2499	TGGACCGTGGACTGGCCGCGCGGTACAAATTTTGGGCACACTTTCAGCTACACAGCCG	2558
Qy	2461	TCTATAATGAGCCGAGAACCTTAATCGCTACTGGACCAACCAACGAGACACTGATTGTG	2520
Db	2559	TCTATAATGAGCCGAGAACCTTAATCGCTACTGGACCAACCAACGAGACACTGATTGTG	2618
Qy	2521	GAGCTGCTGTTTTCAGGGAAGGAACCCGGGTGTGCTCGGAACTCTCCATGCTCGCTTG	2580
Db	2619	GAGCTGCTGTTTTCAGGGAAGGAACCCGGGTGTGCTCGGAACTCTCCATGCTCGCTTG	2678
Qy	2581	GGGACCGAGAGACGCCCTTCGCCAGCCAGCTACACTTGGGCCATCGTGGCTCTGAG	2640
Db	2679	GGGACCGAGAGAGACGCCCTTCGCCAGCCAGCTACACTTGGGCCATCGTGGCTCTGAG	2738
Qy	2641	TGCTCCGTGTCTCGGAGGGGGACAGATGACCGTGAGAGAGGCTGTACAGAGACTTG	2700
Db	2739	TGCTCCGTGTCTCGGAGGGGGACAGATGACCGTGAGAGAGGCTGTACAGAGACTTG	2798
Qy	2701	AAGTTTCAAGTAAATATGTCTTCTGCAATCCCAAGACAGCACTGTACAGGGCTGTGTG	2760
Db	2799	AAGTTTCAAGTAAATATGTCTTCTGCAATCCCAAGACAGCACTGTACAGGGCTGTGTG	2858
Qy	2761	CCTTGCAGAGTATCTGCTGTCTCTCCAGCTGGTCCGTGSGGAACTCGAGTGCCTGAGT	2820
Db	2859	CCTTGCAGAGTATCTGCTGTCTCTCCAGCTGGTCCGTGSGGAACTCGAGTGCCTGAGT	2918
Qy	2821	CGGACGTGTGGGGGGTGCACAGAGCCGCCCGTGCAGTGCACAGCGGGGTGCATAT	2880

Db 2919 CGGACGTGTGGCGGGGGTGGCCAGAGCGCGCCCGTGCAGTGCACACGCGGGTGCACATAT 2978
QY 2881 GACTCGGAGCAAGTCCCGGCGACCCCTGTGCCCCCTCAGCCTGCTCCCTCCAGCAGCGCAGGCC 2940
Db 2979 GACTCGGAGCAAGTCCCGGCGACCCCTGTGCCCCCTCAGCCTGCTCCCTCCAGCAGCGCAGGCC 3038
QY 2941 TGCACCTCTCAGAGCTGCCACCTGTGATGAGAGCGCGGGCCCTGGGCGAGTGCTCACAC 3000
Db 3039 TGCACCTCTCAGAGCTGCCACCTGTGATGAGAGCGCGGGCCCTGGGCGAGTGCTCACAC 3098
QY 3001 ACCTGTGGGAAGGGGTGGAGGAAGCGGCGAGTGGCTGTAAAGAGCAACACCCCTCGGCC 3060
Db 3099 ACCTGTGGGAAGGGGTGGAGGAAGCGGCGAGTGGCTGTAAAGAGCAACACCCCTCGGCC 3158
QY 3061 AGAGCGCAGCTGTGCTGCCCGACGCTGTCTGCACCTCCGAGCCCAAGCCCAAGGATGCATGAA 3120
Db 3159 AGAGCGCAGCTGTGCTGCCCGACGCTGTCTGCACCTCCGAGCCCAAGCCCAAGGATGCATGAA 3218
QY 3121 GCTGTGCTGTTAGCGCTGCGCACAAAGCCCAAGAGCTGAGTGCTGTGCTGCTCGGCTGG 3180
Db 3219 GCTGTGCTGTTAGCGCTGCGCACAAAGCCCAAGAGCTGAGTGCTGTGCTGCTCGGCTGG 3278
QY 3181 TCCAGTGTCTGTGACATGTGAAGAGGAGCAACAGAAAGATTTCTTAAATGTGCTGAA 3240
Db 3279 TCCAGTGTCTGTGACATGTGAAGAGGAGCAACAGAAAGATTTCTTAAATGTGCTGAA 3338
QY 3241 AAGTATGTTTCTGGAAAGTATCGAGAGCTGGCTCAAGAAAGTGTCTCAATTTGCCGAAG 3300
Db 3339 AAGTATGTTTCTGGAAAGTATCGAGAGCTGGCTCAAGAAAGTGTCTCAATTTGCCGAAG 3398
QY 3301 CCCAGCTCGAGTGAACGCTGCTGCGCCCGCTTCCATGCCCCAGGCAACCCCAATTT 3360
Db 3399 CCCAGCTCGAGTGAACGCTGCTGCGCCCGCTTCCATGCCCCAGGCAACCCCAATTT 3458
QY 3361 GCTGTGCGGGAGCCCTCGAGGGGCGAGCTGTTTGGCTCACCCTGCTCAGTGCGACGGCC 3420
Db 3459 GCTGTGCGGGAGCCCTCGAGGGGCGAGCTGTTTGGCTCACCCTGCTCAGTGCGACGGCC 3518
QY 3421 AGCTGTGGGGAGGCGCTTCAGACGAGTCCGTGCACTGCTGCTGGGGGCGCGCGGCC 3480
Db 3519 AGCTGTGGGGAGGCGCTTCAGACGAGTCCGTGCACTGCTGCTGGGGGCGCGCGGCC 3578
QY 3481 TCAGGTGCTCTCTGACACAGAGCCTTCGGCTTCCCTGGCTGCGAACACTCACTTCTGC 3540
Db 3579 TCAGGTGCTCTCTGACACAGAGCCTTCGGCTTCCCTGGCTGCGAACACTCACTTCTGC 3638
QY 3541 CCCATTGCAGAGAAAGATGCTTCTGCAAGACTACTTCCACTGGTGCTACCTGTA 3600
Db 3639 CCCATTGCAGAGAAAGATGCTTCTGCAAGACTACTTCCACTGGTGCTACCTGTA 3698
QY 3601 CCCAGCAGCGGATGTGACAGCCCAAGTTCTACGGCAAGCAGTGCTGCAAGACTTGTCT 3660
Db 3699 CCCAGCAGCGGATGTGACAGCCCAAGTTCTACGGCAAGCAGTGCTGCAAGACTTGTCT 3758
QY 3661 AAGTCCAACTTGTGA 3675
Db 3759 AAGTCCAACTTGTGA 3773

RESULT 5

US-10-296-616-1
; Sequence 1, Application US/10296616
; Publication No. US20030129658A1
; GENERAL INFORMATION:
; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: No. US20030129658A1e1 protease
; FILE REFERENCE: Y0132PCT-664
; CURRENT APPLICATION NUMBER: US/10/296,616
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: JP 2000-393372
; PRIOR FILING DATE: 2000-12-25
; NUMBER OF SEQ ID NOS: 26

; SEQ ID NO 1
; LENGTH: 3675
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3675)
US-10-296-616-1

Query Match 100.0%; Score 3673.4; DB 15; Length 3675;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3674; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAAGCCCCCGCGCGCGATGGCGGGCTTGGCGCGCTGTGGATGCTGTGGCGCAG 60
Db 1 ATGAAGCCCCCGCGCGCGATGGCGGGCTTGGCGCGCTGTGGATGCTGTGGCGCAG 60

QY 61 GTGGCGGAGCAGGCACTTGGCTGGCCATGGGACCCGCGAGCGCGCGCGCTGGGAGCCCG 120
Db 61 GTGGCGGAGCAGGCACTTGGCTGGCCATGGGACCCGCGAGCGCGCGCGCTGGGAGCCCG 120

QY 121 AGCGTCCCGCGCTCTCTCCACCCGCGAGCGCGCGGCTGGATGGAAGGCGCAATAT 180
Db 121 AGCGTCCCGCGCTCTCTCCACCCGCGAGCGCGCGGCTGGATGGAAGGCGCAATAT 180

QY 181 GACCTGGTCTCTGCTACGAGTTGAACACAGGGGCGATTACGTGTCCATGAAATCATG 240
Db 181 GACCTGGTCTCTGCTACGAGTTGAACACAGGGGCGATTACGTGTCCATGAAATCATG 240

QY 241 CACCATCAGCGCGGAGAGAGCAGTGGCGGTGTCGAGTTGAGTCTCTTCACTTCG 300
Db 241 CACCATCAGCGCGGAGAGAGCAGTGGCGGTGTCGAGTTGAGTCTCTTCACTTCG 300

QY 301 CTGAAAGGCTCCAGGCAACGACTTCCACGTGGATCTCAGGACTTCCAGCAGCCTAGTGGCT 360
Db 301 CTGAAAGGCTCCAGGCAACGACTTCCACGTGGATCTCAGGACTTCCAGCAGCCTAGTGGCT 360

QY 361 CTTGGCTTTATTTGTGAGACGTTGGAAAGACAGGCACCTAACTCTGTGACAGCTTTACCG 420
Db 361 CTTGGCTTTATTTGTGAGACGTTGGAAAGACAGGCACCTAACTCTGTGACAGCTTTACCG 420

QY 421 CCAGAGGACTTCTGTTCTATCAAGGCTCTTTGCGATCACACAGAACTCTCAGTGGCC 480
Db 421 CCAGAGGACTTCTGTTCTATCAAGGCTCTTTGCGATCACACAGAACTCTCAGTGGCC 480

QY 481 CTTTCAACTGCAAGGCTTGTGAGCATGATACGAAAGAGAGGAGGAGTACTTCTCTA 540
Db 481 CTTTCAACTGCAAGGCTTGTGAGCATGATACGAAAGAGAGGAGGAGTACTTCTCTA 540

QY 541 AGGCCACTTCTTCAACCTCTCATGGAACTCTGGCAGAGCTGCCAAGGCGAGCTCGCCA 600
Db 541 AGGCCACTTCTTCAACCTCTCATGGAACTCTGGCAGAGCTGCCAAGGCGAGCTCGCCA 600

QY 601 TCCACAGTACTGTACAGAGATCCACAGAGCCCACTGCTCTCTGGGGCCAGTGAAGTCTG 660
Db 601 TCCACAGTACTGTACAGAGATCCACAGAGCCCACTGCTCTCTGGGGCCAGTGAAGTCTG 660

QY 661 GTGACCTCAAGGACATGGAGCTGGCAATCAACCCCTGCAAGAGGAGGAGTCTGGCCCTG 720
Db 661 GTGACCTCAAGGACATGGAGCTGGCAATCAACCCCTGCAAGAGGAGGAGTCTGGCCCTG 720

QY 721 GGACTGCCACAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCCAGCCTCCC 780
Db 721 GGACTGCCACAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCCAGCCTCCC 780

QY 781 AAGGAAGACCTTCTTCACTTGGCAGATGAGTATAGTCTTGTAGGATAGCGCTCT 840
Db 781 AAGGAAGACCTTCTTCACTTGGCAGATGAGTATAGTCTTGTAGGATAGCGCTCT 840

QY 841 CTTCTGAGGTCCCATAGAAATGAAGAACTGAACGTTGGAGACCTTGTGTGTGTCGACAAA 900
Db 841 CTTCTGAGGTCCCATAGAAATGAAGAACTGAACGTTGGAGACCTTGTGTGTGTCGACAAA 900

Qy	901	AAGATGATGCAAAACCATGGCCATGAAATATCACCACTAGTGTCTCAGTACTCAAC	960
Db	901	AAGATGATGCAAAACCATGGCCATGAAATATCACCACTAGTGTCTCAGTACTCAAC	960
Qy	961	ATGCTATCTGCTTTATTTCAAGATGGAAACAATAGGAGGAAACATCAACATTTGCAATTGTA	1020
Db	961	ATGCTATCTGCTTTATTTCAAGATGGAAACAATAGGAGGAAACATCAACATTTGCAATTGTA	1020
Qy	1021	GGTCTGATTTCTTTAGAAGATGAACAGCAGACCTGGTGATAGTCAACGACGACACAC	1080
Db	1021	GGTCTGATTTCTTTAGAAGATGAACAGCAGACCTGGTGATAGTCAACGACGACACAC	1080
Qy	1081	ACCTTAAGTAGTCTTTCGACGCTGGTCTGGATTTGATGGGGAAGATGGGACTCGTCAT	1140
Db	1081	ACCTTAAGTAGTCTTTCGACGCTGGTCTGGATTTGATGGGGAAGATGGGACTCGTCAT	1140
Qy	1141	GACCAGCCCATCTTACTGACTGGTCTGGATATGTTCTCTGGAAAGATGGGACTCGTCAT	1200
Db	1141	GACCAGCCCATCTTACTGACTGGTCTGGATATGTTCTCTGGAAAGATGGGACTCGTCAT	1200
Qy	1201	ACTTTGGGATTTGCAACCCATAAGTGGAAATGTAGTAAATATCCGAGCTGCACGATTAAT	1260
Db	1201	ACTTTGGGATTTGCAACCCATAAGTGGAAATGTAGTAAATATCCGAGCTGCACGATTAAT	1260
Qy	1261	GAAGATACAGGCTTTGGACTGGCCCTTACCAATGCCCATGGCTGAGTCTGGAACAACATTTGGC	1320
Db	1261	GAAGATACAGGCTTTGGACTGGCCCTTACCAATGCCCATGGCTGAGTCTGGAACAACATTTGGC	1320
Qy	1321	ATGATTCATGATGGAGAAAGGAAACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCT	1380
Db	1321	ATGATTCATGATGGAGAAAGGAAACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCT	1380
Qy	1381	ACATTTGGCAGGACGCAATGGAGTCTTCTCTGGTCACTTGATGCAAGAGCTGTGAAGAA	1440
Db	1381	ACATTTGGCAGGACGCAATGGAGTCTTCTCTGGTCACTTGATGCAAGAGCTGTGAAGAA	1440
Qy	1441	AAATTTCTAAGACCGCTCAAGCTATCTGCTTGCTGATCAGCAAAAGCTGTGAAGAA	1500
Db	1441	AAATTTCTAAGACCGCTCAAGCTATCTGCTTGCTGATCAGCAAAAGCTGTGAAGAA	1500
Qy	1501	TACAAGTATCTTGAGAAATTTGCCAGGAAATATATGATGCAAAACACACAGTGCAGTGG	1560
Db	1501	TACAAGTATCTTGAGAAATTTGCCAGGAAATATATGATGCAAAACACACAGTGCAGTGG	1560
Qy	1561	CAGTTCGGAGAGAAAGCCAAAGCTCTGCATGCTGGACTTTAAAGAGACATCTGTAAAGCC	1620
Db	1561	CAGTTCGGAGAGAAAGCCAAAGCTCTGCATGCTGGACTTTAAAGAGACATCTGTAAAGCC	1620
Qy	1621	CTGTGGTGGCATCGTATTTGGAAGAAATGTGAGACTTAAATTTATGCCAGCAGAGAAGC	1680
Db	1621	CTGTGGTGGCATCGTATTTGGAAGAAATGTGAGACTTAAATTTATGCCAGCAGAGAAGC	1680
Qy	1681	ACAAATTTGTGGGCAATGACATGTTGGTCCGGGAGGACAGTGTGTGAATATGGTGATGA	1740
Db	1681	ACAAATTTGTGGGCAATGACATGTTGGTCCGGGAGGACAGTGTGTGAATATGGTGATGA	1740
Qy	1741	GGCCCCAAGCCCACTGGCCACTGGTTCGACTGGTCTTCTGGTCCCCATGCTCCAGG	1800
Db	1741	GGCCCCAAGCCCACTGGCCACTGGTTCGACTGGTCTTCTGGTCCCCATGCTCCAGG	1800
Qy	1801	ACTTCGGAGGGGAGTATCTCATAGGAGTGGCCCTCTGCACCAACCCCAAGCCATCGCAT	1860
Db	1801	ACTTCGGAGGGGAGTATCTCATAGGAGTGGCCCTCTGCACCAACCCCAAGCCATCGCAT	1860
Qy	1861	GGAGGGAAGTCTCTGAGGGCTCCACTCGCATCTGGAAGCTCTGCAACAGTTCAGAAATGT	1920
Db	1861	GGAGGGAAGTCTCTGAGGGCTCCACTCGCATCTGGAAGCTCTGCAACAGTTCAGAAATGT	1920
Qy	1921	CCCCGGGACAGTGTGAATCTCCGTGCTGCTCAGTGTGCCGAGCACAACGACGAGATTC	1980
Db	1921	CCCCGGGACAGTGTGAATCTCCGTGCTGCTCAGTGTGCCGAGCACAACGACGAGATTC	1980
Qy	1981	AGAGGGCGGCACCTACAAGTGGAAAGCCCTTACCTCAAGTAGAAGATCAGGACTTATGCAAA	2040

[illegible]

Db 3061 AGAGCGAGCTGCTGCCGAGCGCTGTGTGCACTCCGAGCCCAAGCCAGGATGCAATGAA 3120
Qy 3121 GCCTGTCTGCTTCAGCGCTGCCCAAGCCCAAGAGCTGCAGTGGCTGTGTGCCGCTGG 3180
Db 3121 GCCTGTCTGCTTCAGCGCTGCCCAAGCCCAAGAGCTGCAGTGGCTGTGTGCCGCTGG 3180
Qy 3181 TCCAGTGTCTGTGATGTGAAGAGGAGACACAGAAAGATCTTAAATGTGCTGAA 3240
Db 3181 TCCAGTGTCTGTGATGTGAAGAGGAGACACAGAAAGATCTTAAATGTGCTGAA 3240
Qy 3241 AGATATGTTTCTGAAAGTATCCAGAGCTGGCTCAAGAGAGTGTCCATTTTGCAGAG 3300
Db 3241 AGATATGTTTCTGAAAGTATCCAGAGCTGGCTCAAGAGAGTGTCCATTTTGCAGAG 3300
Qy 3301 CCAGCTGTGAGCTGAAACGTGCTGCCGCCCGCTTCCATGCCCGGACACCCCAATTT 3360
Db 3301 CCAGCTGTGAGCTGAAACGTGCTGCCGCCCGCTTCCATGCCCGGACACCCCAATTT 3360
Qy 3361 GCTGCTGCGGACCTTCGAGGGGACGCTGGTTTGGCTCACCCTGGTCTCAGTGCACGGCC 3420
Db 3361 GCTGCTGCGGACCTTCGAGGGGACGCTGGTTTGGCTCACCCTGGTCTCAGTGCACGGCC 3420
Qy 3421 AGCTGTGGGGAGCGCTTCAGACGAGTCCGTGCACTGCTGGCTGGGGCGCGCGCGCC 3480
Db 3421 AGCTGTGGGGAGCGCTTCAGACGAGTCCGTGCACTGCTGGCTGGGGCGCGCGCGCC 3480
Qy 3481 TCAGGCTGCTCTGCACAGAGCCTTCGGCTCCCTGCGCTGCAACACTCACTTTCTGC 3540
Db 3481 TCAGGCTGCTCTGCACAGAGCCTTCGGCTCCCTGCGCTGCAACACTCACTTTCTGC 3540
Qy 3541 CCATTTGCAGAGAGAGATGCTCTTCGAAAGACTTTCACACTGCTGGTCTACTTGGTA 3600
Db 3541 CCATTTGCAGAGAGAGATGCTCTTCGAAAGACTTTCACACTGCTGGTCTACTTGGTA 3600
Qy 3601 CCCACGCGGATGTCAGCCACAGTTCAGGCAAGCTGCGCAAGCAGTGTGCAAGCTTCTCT 3660
Db 3601 CCCACGCGGATGTCAGCCACAGTTCAGGCAAGCTGCGCAAGCAGTGTGCAAGCTTCTCT 3660
Qy 3661 AAGTCCAACTTGTGA 3675
Db 3661 AAGTCCAACTTGTGA 3675

RESULT 6

US-10-330-176-1
; Sequence 1, Application US/10330176
; Publication No. US20030228676A1
; GENERAL INFORMATION:
; APPLICANT: Agostino, Michael
; APPLICANT: DiBlasio, Elizabeth
; TITLE OF INVENTION: AGGREGANASE MOLECULES
; FILE REFERENCE: AM100884
; CURRENT APPLICATION NUMBER: US/10/330,176
; PRIOR FILING DATE: 2002-12-30
; PRIOR APPLICATION NUMBER: 60/344,895
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 3675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-330-176-1

Query Match 99.9%; Score 3671.8; DB 17; Length 3675;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3673; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGAAGCCCGCGCGCGATGCGGGGCTTGGCGGCTGTGGATGCTGTGGCGCAG 60
Db 1 ATGAAGCCCGCGCGCGATGCGGGGCTTGGCGGCTGTGGATGCTGTGGCGCAG 60
Qy 61 GTGGCGGAGCAGGCACTTGTGTGCGCCATGGGACCGCGAGCGGCTGGGAGCCCG 120

Db 61 GTGGCGGAGCAGGCACTTGTGTGCGCCATGGGACCGCGAGCGGCTGGGAGCCCG 120
Qy 121 AGCGTCCCGCGTCTCTCTCCACCCCGGAGCGCGCGGCTGGATGAAAAGGCGCAATAT 180
Db 121 AGCGTCCCGCGTCTCTCTCCACCCCGGAGCGCGCGGCTGGATGAAAAGGCGCAATAT 180
Qy 181 GACCTGGTCTCTGCTTACCGAGTTGACCAAGGGGCGATTAACGTTCCCAATGAATCATG 240
Db 181 GACCTGGTCTCTGCTTACCGAGTTGACCAAGGGGCGATTAACGTTCCCAATGAATCATG 240
Qy 241 CACCATCAGGGGCGGAGAGAGCGAGTGGCGGTGTCGAGGTTGAGTCTCTTCACTTCGG 300
Db 241 CACCATCAGGGGCGGAGAGAGCGAGTGGCGGTGTCGAGGTTGAGTCTCTTCACTTCGG 300
Qy 301 CTGAAAGGCTTCCAGGCAACGACTTCCACGTTGATCTGAGGACTTCCAGAGCGCTAGTGGCT 360
Db 301 CTGAAAGGCTTCCAGGCAACGACTTCCACGTTGATCTGAGGACTTCCAGAGCGCTAGTGGCT 360
Qy 361 CTTGGCTTTATTTGTGAGAGCTTTGGGAAAGACAGGCACTAAAGTCTGTGCAAGCTTACCG 420
Db 361 CTTGGCTTTATTTGTGAGAGCTTTGGGAAAGACAGGCACTAAAGTCTGTGCAAGCTTACCG 420
Qy 421 CCAGAGGACTTCTGTCTTCTATCAAGGCTCTTTGCGATCAACAGAACTCTCTCAGTGGCC 480
Db 421 CCAGAGGACTTCTGTCTTCTATCAAGGCTCTTTGCGATCAACAGAACTCTCTCAGTGGCC 480
Qy 481 CTTTCAACCTGCGCAAGGCTTGTGAGGATGATACGAAACAGAGAGGCGAGATTAATTCCTA 540
Db 481 CTTTCAACCTGCGCAAGGCTTGTGAGGATGATACGAAACAGAGAGGCGAGATTAATTCCTA 540
Qy 541 AGGCCACTTCTCTTCAACCTCTCATGGAACCTCGGAGAGCTGCCCAAGGCGCTCGCCA 600
Db 541 AGGCCACTTCTCTTCAACCTCTCATGGAACCTCGGAGAGCTGCCCAAGGCGCTCGCCA 600
Qy 601 TCCAGCTACTGTACAGAGATCCACAGCGCCCATGCTCTCGGGGCGAGTGAAGTCTCTG 660
Db 601 TCCAGCTACTGTACAGAGATCCACAGCGCCCATGCTCTCGGGGCGAGTGAAGTCTCTG 660
Qy 661 GTGACCTCAAGGACATGCGAGCTGGCAGATCAACCTCGCACAGAGCGAGCTTCCGCTG 720
Db 661 GTGACCTCAAGGACATGCGAGCTGGCAGATCAACCTCGCACAGAGCGAGCTTCCGCTG 720
Qy 721 GGAGTCCCAAAAGCAGCAATTTCTGTGGAAGACGCAAGAAATACATGCCAGCTTCCC 780
Db 721 GGAGTCCCAAAAGCAGCAATTTCTGTGGAAGACGCAAGAAATACATGCCAGCTTCCC 780
Qy 781 AAGGAAGACCTCTTCACTCTTGGCAGATGATTAAGTCTTGTACGGCATTAAGCGCTCT 840
Db 781 AAGGAAGACCTCTTCACTCTTGGCAGATGATTAAGTCTTGTACGGCATTAAGCGCTCT 840
Qy 841 CTTCTGAGGTCCCATAGAAATGAAGAACTGAAAGTGGAGACCTTGGTGGTGGTGGACAAA 900
Db 841 CTTCTGAGGTCCCATAGAAATGAAGAACTGAAAGTGGAGACCTTGGTGGTGGTGGACAAA 900
Qy 901 AAGATGATGCAAAAACCATGCGCCATGAAAATATCACCACTACGTGTCTCAGTACTCAAC 960
Db 901 AAGATGATGCAAAAACCATGCGCCATGAAAATATCACCACTACGTGTCTCAGTACTCAAC 960
Qy 961 ATGGTATCTGCTTTTATTAAGATGGAACAATAGGAGGAAACATCAACATTTGTA 1020
Db 961 ATGGTATCTGCTTTTATTAAGATGGAACAATAGGAGGAAACATCAACATTTGTA 1020
Qy 1021 GGTCTGATTTCTTAGAGATGAACAGCCAGGACTGGTATTAAGTCAACCGCAGACCCAC 1080
Db 1021 GGTCTGATTTCTTAGAGATGAACAGCCAGGACTGGTATTAAGTCAACCGCAGACCCAC 1080
Qy 1081 ACCTTAAGTAGTCTTCTGCGAGTGGCAGTCTGGAATGATGGGAAAGATGGGACTCTCAT 1140
Db 1081 ACCTTAAGTAGTCTTCTGCGAGTGGCAGTCTGGAATGATGGGAAAGATGGGACTCTCAT 1140
Qy 1141 GACCGAGCCATCTTACTGATCTGCTGTGGATATATGTTCTCTGGAAGATGAGCCTCTGAC 1200
Db 1141 GACCGAGCCATCTTACTGATCTGCTGTGGATATATGTTCTCTGGAAGATGAGCCTCTGAC 1200

D	b	1141	GACCA CGCCATCTTACTGACTGGTCTGGATATATGTTCCTGGAAAGATGAGCCCTCTGCAC	1200
Q	y	1201	ACTTTGGGATTGTGCCACCACCAATAAGTGAATGTAGTAATAATACGACCTGCACGATTAAAT	1260
D	b	1201	ACTTTGGGATTGTGCCACCACCAATAAGTGAATGTAGTAATAATACGACCTGCACGATTAAAT	1260
Q	y	1261	GAAGATACAGGCTTTGGACTGGCCCTTCACCATTCGCCATGAGTCTGGACACAACCTTTGGC	1320
D	b	1261	GAAGATACAGGCTTTGGACTGGCCCTTCACCATTCGCCATGAGTCTGGACACAACCTTTGGC	1320
Q	y	1321	ATGATTTCATGATGAGAGAAGGGAACAATGTGTAAAAGTCCGAGGGGCAACATCATGTCCCCT	1380
D	b	1321	ATGATTTCATGATGAGAGAAGGGAACAATGTGTAAAAGTCCGAGGGGCAACATCATGTCCCCT	1380
Q	y	1381	ACATTGGCAGGACGCAATGGAGTCTTCTCTGTGCTCACCTTCGACCGCCGACGATCTACAC	1440
D	b	1381	ACATTGGCAGGACGCAATGGAGTCTTCTCTGTGCTCACCTTCGACCGCCGACGATCTACAC	1440
Q	y	1441	AAATTTCTTAAGCACCGCTCAAGCTATCTGCTTGCTGCTGATCAGCCAAGGCTGTGAAGGAA	1500
D	b	1441	AAATTTCTTAAGCACCGCTCAAGCTATCTGCTTGCTGCTGATCAGCCAAGGCTGTGAAGGAA	1500
Q	y	1501	TACAAAGTATCTTGAGAAATTCGACAGGAGAATTATATGATGCAAAACACACAGTGCAGGTGG	1560
D	b	1501	TACAAAGTATCTTGAGAAATTCGACAGGAGAATTATATGATGCAAAACACACAGTGCAGGTGG	1560
Q	y	1561	CAGTTTCGAGAGAAGCCAGCTCTGATCTGCTGGAATTTAAAAAGGACATCTGTAAAGCC	1620
D	b	1561	CAGTTTCGAGAGAAGCCAGCTCTGATCTGCTGGAATTTAAAAAGGACATCTGTAAAGCC	1620
Q	y	1621	CTGTGGTGGCCATCGTATTGGAAAGAAATGTGAGACTTAAATTTATGCCAGCAGCAGAGGC	1680
D	b	1621	CTGTGGTGGCCATCGTATTGGAAAGAAATGTGAGACTTAAATTTATGCCAGCAGCAGAGGC	1680
Q	y	1681	ACAAATTTGTGGGCATGACATGTGGTCCGAGGAGGACAGTGTGTAATAATGGTAGTAA	1740
D	b	1681	ACAAATTTGTGGGCATGACATGTGGTCCGAGGAGGACAGTGTGTAATAATGGTAGTAA	1740
Q	y	1741	GGCCCCAAGGCCAACCATGSCCACTGGTCCGACTGGTCTTCTGTTGGTCCCAATGCTCCAGG	1800
D	b	1741	GGCCCCAAGGCCAACCATGSCCACTGGTCCGACTGGTCTTCTGTTGGTCCCAATGCTCCAGG	1800
Q	y	1801	ACCTGCGGAGGGGGAGTATCTCATAGGATGCGCTCTGCACCAACCCCAAGCCATCGCAT	1860
D	b	1801	ACCTGCGGAGGGGGAGTATCTCATAGGATGCGCTCTGCACCAACCCCAAGCCATCGCAT	1860
Q	y	1861	GGAGGGAAGTTCTGTGAGGGCTCCACTCGCACTCTGAAAGCTCTGAAACAGTCTGAAATGT	1920
D	b	1861	GGAGGGAAGTTCTGTGAGGGCTCCACTCGCACTCTGAAAGCTCTGAAACAGTCTGAAATGT	1920
Q	y	1921	CCCCGGGACAGTGTGACTTCCGTGCTGCTCAGTGTGCGGAGCACAACAGCAGACCGATTTC	1980
D	b	1921	CCCCGGGACAGTGTGACTTCCGTGCTGCTCAGTGTGCGGAGCACAACAGCAGACCGATTTC	1980
Q	y	1981	AGAGGGCGGCACTACAAGTGGAGGCTTACACTCAAGTAGAAGATCAGGACTTATGCAAA	2040
D	b	1981	AGAGGGCGGCACTACAANGTGGAGGCTTACACTCAAGTAGAAGATCAGGACTTATGCAAA	2040
Q	y	2041	CTCTACTGTATCGCAGAAGATTTGATTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	2100
D	b	2041	CTCTACTGTATCGCAGAAGATTTGATTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	2100
Q	y	2101	GGGACTCCATGCTCCGAGGATAGCCGTAATGTTTGTTATAGATGGATATGTGAGAGATT	2160
D	b	2101	GGGACTCCATGCTCCGAGGATAGCCGTAATGTTTGTTATAGATGGGATATGTGAGAGATT	2160
Q	y	2161	GGATGTGA CAATGTCTTGGATCTGATGCTGTGTGAAGACCTGTGTGGGGTGTGTACGGG	2220
D	b	2161	GGATGTGA CAATGTCTTGGATCTGATGCTGTGTGAAGACCTGTGTGGGGTGTGTACGGG	2220
Q	y	2221	AATAACTCAGCTGCACGATTTCACAGGGTCTCTACACCAAGCACCAACCAACCAAG	2280
D	b	2221	AATAACTCAGCTGCACGATTTCACAGGGTCTCTACACCAAGCACCAACCAACCAAG	2280

QY 3361 GCTGCTGCGGACCCCTCGAGGGGACAGTGGTTTGCTCACCCTGGTCTCAGTGCACGGCC 3420
DB 3361 GCTGCTGCGGACCCCTCGAGGGGACAGTGGTTTGCTCACCCTGGTCTCAGTGCACGGCC 3420
QY 3421 AGCTGTGGGGAGGCGCTTCAGACGAGGTCCGTGCAGTGGCTGGCTGGGGGCGGCGGCG 3480
DB 3421 AGCTGTGGGGAGGCGCTTCAGACGAGGTCCGTGCAGTGGCTGGCTGGGGGCGGCGGCG 3480
QY 3481 TCAGGCTGCTCTCTGCACACAGAGCCTTCGGCCCTCCCTGGCCCTGCAACAACCTCACTTCTGC 3540
DB 3481 TCAGGCTGCTCTCTGCACACAGAGCCTTCGGCCCTCCCTGGCCCTGCAACAACCTCACTTCTGC 3540
QY 3541 CCCATTGCAAGAGAAAGTCTCTGCAAGACTACTTCCACTGGTCTACTCTGTA 3600
DB 3541 CCCATTGCAAGAGAAAGTCTCTGCAAGACTACTTCCACTGGTCTACTCTGTA 3600
QY 3601 CCCACGACGGGATGTCAGCCACAAGTTCACGGCAAGCAGTGTGCAAGACTTCTCT 3660
DB 3601 CCCACGACGGGATGTCAGCCACAAGTTCACGGCAAGCAGTGTGCAAGACTTCTCT 3660
QY 3661 AAGTCCAACTTGTGA 3675
DB 3661 AAGTCCAACTTGTGA 3675

RESULT 7

US-10-275-107-16
; Sequence 16, Application US/10275107
; Publication No. US20040063107A1
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY D.
; APPLICANT: WHYTE, DAVID
; APPLICANT: SUDARSANAM, SUCHA
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEEL, SEAN R.
; APPLICANT: PAYNE, VILJA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1479
; CURRENT APPLICATION NUMBER: US/10/275,107
; CURRENT FILING DATE: 2003-11-03
; PRIOR APPLICATION NUMBER: PCT/US01/14431
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/201,879
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 3675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-275-107-16

Query Match 99.9%; Score 3671.8; DB 18; Length 3675;
Best Local Similarity 99.9%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;
Matches 3673; Conservative 0;
QY 1 ATGAAGCCCGCGCGCGGATGGCGGGCTTGGCGCGCTGTGGATGCTGTGGCGCAG 60
DB 1 ATGAAGCCCGCGCGCGGATGGCGGGCTTGGCGCGCTGTGGATGCTGTGGCGCAG 60
QY 61 GTGGCCGAGGAGCAGTGGTGGCGCCATGGACCCGCGAGCGGCGCGCTGGAGCCCG 120
DB 61 GTGGCCGAGGAGCAGTGGTGGCGCCATGGACCCGCGAGCGGCGCGCTGGAGCCCG 120
QY 121 AGCGTCCCGCTCTCTCCACCGCGGAGCGCGGGCTGGATGGAAGCGGCAATAT 180
DB 121 AGCGTCCCGCTCTCTCTCCACCGCGGAGCGCGGGCTGGATGGAAGCGGCAATAT 180
QY 181 GACCTGCTCTCTGCTACGAGGTGTGACACAGGGGCGATTAAGTGTCCCATGAAATCATG 240
DB 181 GACCTGCTCTCTGCTACGAGGTGTGACACAGGGGCGATTAAGTGTCCCATGAAATCATG 240
QY 241 CACCATCAGCGGCGGAGAGAGCAGTGGCGGTGTCCGAGGTTGAGTCTCTTCACTTCCG 300

DB 241 CACCATCAGCGGCGGAGAGAGCAGTGGCGGTGTCCGAGTTGAGTCTCTTCACTTCG 300
QY 301 CTGAAGAGCTCCAGGCA CGACTTCCA CGTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360
DB 301 CTGAAGAGCTCCAGGCA CGACTTCCA CGTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360
QY 361 CTGGCTTTATTTGTGAGAGCTTTGGGAAAGACAGGCACTAAGTCTGTGAGACTTTACCG 420
DB 361 CTGGCTTTATTTGTGAGAGCTTTGGGAAAGACAGGCACTAAGTCTGTGAGACTTTACCG 420
QY 421 CCAGAGGACTCTCTGTTCTTATCAAGGCTCTTTGCGATCACAGAAACTCTCTCAGTGGCC 480
DB 421 CCAGAGGACTCTCTGTTCTTATCAAGGCTCTTTGCGATCACAGAAACTCTCTCAGTGGCC 480
QY 481 CTTTCAACCTGCGCAAGGCTTGT CAGGCATGATACGAA CAGAAAGAGCGAGATTACTTCTTA 540
DB 481 CTTTCAACCTGCGCAAGGCTTGT CAGGCATGATACGAA CAGAAAGAGCGAGATTACTTCTTA 540
QY 541 AGGCCACTTCTCTTCA CACTCTCATGGAATCTCGG CAGAGCTGCC CCAAGCAGCTGCCA 600
DB 541 AGGCCACTTCTCTTCA CACTCTCATGGAATCTCGG CAGAGCTGCC CCAAGCAGCTGCCA 600
QY 601 TCCCACTGCTACAGAGATCC CAGAGCCCCCATGCT CCGGGCCAGTGAAGTCCCTG 660
DB 601 TCCCACTGCTACAGAGATCC CAGAGCCCCCATGCT CCGGGCCAGTGAAGTCCCTG 660
QY 661 GTGACCTCAAGCAGATGGGAGCTGG CACATCAACCCCTG CACAGCAGCAGCCTTCCGCTG 720
DB 661 GTGACCTCAAGCAGATGGGAGCTGG CACATCAACCCCTG CACAGCAGCAGCCTTCCGCTG 720
QY 721 GGACTGCCACAAAGCAGCATTTCTGTGGAAGACG CAAAGAAATACATATGCCCCAGCTCCC 780
DB 721 GGACTGCCACAAAGCAGCATTTCTGTGGAAGACG CAAAGAAATACATATGCCCCAGCTCCC 780
QY 781 AAGGAGAGCTCTTCTCATCTTGCAGATGAGTAA GTCTTGTCTTACGGCATAGCCCTCT 840
DB 781 AAGGAGAGCTCTTCTCATCTTGCAGATGAGTAA GTCTTGTCTTACGGCATAGCCCTCT 840
QY 841 CTTCTGAGGTCCCATAGAAATGAAAGCTGGAACCTGGAGACCTTGGTGGTGGTGGCAGAA 900
DB 841 CTTCTGAGGTCCCATAGAAATGAAAGCTGGAACCTGGAGACCTTGGTGGTGGTGGCAGAA 900
QY 901 AAGATGATGCAAAA CCAATGGCCATGAAAATAT CACCACCTAGCTGTCTACGATCTCAAC 960
DB 901 AAGATGATGCAAAA CCAATGGCCATGAAAATAT CACCACCTAGCTGTCTACGATCTCAAC 960
QY 961 ATGGTATCTGCTTTATTCAAAGATGGNA CAAATAGGAGGAAACATCAACATTCGAATTGTA 1020
DB 961 ATGGTATCTGCTTTATTCAAAGATGGNA CAAATAGGAGGAAACATCAACATTCGAATTGTA 1020
QY 1021 GGTCTGATTTCTTAGAAGATGAA CAGCCAGGACTGGTGATTAAGTCA CCA CGCAGACCCAC 1080
DB 1021 GGTCTGATTTCTTAGAAGATGAA CAGCCAGGACTGGTGATTAAGTCA CCA CGCAGACCCAC 1080
QY 1081 ACCTTAAGTGTCTTCTGCGAGTGGCAGTCTGGA TTGATGGGAAAGATGGGACTGCTCAT 1140
DB 1081 ACCTTAAGTGTCTTCTGCGAGTGGCAGTCTGGA TTGATGGGAAAGATGGGACTGCTCAT 1140
QY 1141 GACCAAGCCATCTTACTGATCTGGATATATGTTCTCTGGAAGATGAGCCCTGTGAC 1200
DB 1141 GACCAAGCCATCTTACTGATCTGGATATATGTTCTCTGGAAGATGAGCCCTGTGAC 1200
QY 1201 ACTTTGGGATTTGCA CCCCATAAGTGGAAATGTGTAGTAAATATCGCAGCTGCAAGTAAAT 1260
DB 1201 ACTTTGGGATTTGCA CCCCATAAGTGGAAATGTGTAGTAAATATCGCAGCTGCAAGTAAAT 1260
QY 1261 GAAGATACAGGTCTTGGACTGCGCTTCA CCAATTTGCCCATGAGTCTGGACACAACTTTGCG 1320
DB 1261 GAAGATACAGGTCTTGGACTGCGCTTCA CCAATTTGCCCATGAGTCTGGACACAACTTTGCG 1320
QY 1321 ATGATTATGATGGAGAGGGGAAACATGTGTAA AAGTCCGAGGGGCAACATCATGTCCTCT 1380

Db 1321 ATGATTTCATGAGAGAGGAAACATGTGTGTAAGAGTCCGAGGGCAACATCATGTCCCT 1380
Qy 1381 ACAATGGCAGGACGCAATGGAGTCTTCTCTGTGTCACTTGTGAGCGGCGAGTATCTACAC 1440
Db 1381 ACAATGGCAGGACGCAATGGAGTCTTCTCTGTGTCACTTGTGAGCGGCGAGTATCTACAC 1440
Qy 1441 AATTTCTTAAGCACCCTCAAGCTATCTGCTTGTGATCAGCAAGCCTGTGAAGAA 1500
Db 1441 AATTTCTTAAGCACCCTCAAGCTATCTGCTTGTGATCAGCAAGCCTGTGAAGAA 1500
Qy 1501 TACAAGTATCTGAGAAATGCGAGGAAATATATGATGCAAAACACAGTGAAGTGG 1560
Db 1501 TACAAGTATCTGAGAAATGCGAGGAAATATATGATGCAAAACACAGTGAAGTGG 1560
Qy 1561 CAGTTCCGAGAGAAAGCCAGCTCTGATGCTGGACTTTTAAAGAGCATCTGTAAAGCC 1620
Db 1561 CAGTTCCGAGAGAAAGCCAGCTCTGATGCTGGACTTTTAAAGAGCATCTGTAAAGCC 1620
Qy 1621 CTGTGTGCCATCGTATGGAAGAAATGTGAGACTTAAATTTATGCCAGCAGCAAGGC 1680
Db 1621 CTGTGTGCCATCGTATGGAAGAAATGTGAGACTTAAATTTATGCCAGCAGCAAGGC 1680
Qy 1681 ACAATTTGTGGCATGACATGTGGTCCGAGGAGACAGTGTGAAATATGTTGATGAA 1740
Db 1681 ACAATTTGTGGCATGACATGTGGTCCGAGGAGACAGTGTGAAATATGTTGATGAA 1740
Qy 1741 GGGCCCAAGCCCAACCATGCGCACTGGTCCGACTGGTCTTCTGCTCCCATGCTCCAGG 1800
Db 1741 GGGCCCAAGCCCAACCATGCGCACTGGTCCGACTGGTCTTCTGCTCCCATGCTCCAGG 1800
Qy 1801 ACTCGGAGGGGAGTATCTCATAGGAGTGGCTCTGCAACACCCCAAGCCATCGCAT 1860
Db 1801 ACTCGGAGGGGAGTATCTCATAGGAGTGGCTCTGCAACACCCCAAGCCATCGCAT 1860
Qy 1861 GAGGGAGATCTGTGAAGGCTCACTGCACTCTGAAGTCTGCAACAGTCAAGAAATGT 1920
Db 1861 GAGGGAGATCTGTGAAGGCTCACTGCACTCTGAAGTCTGCAACAGTCAAGAAATGT 1920
Qy 1921 CCCCGGAGACGTGTGTGACTTCCGCTGCTCAGTGTCCGAGCAACACGACGAGATTC 1980
Db 1921 CCCCGGAGACGTGTGTGACTTCCGCTGCTCAGTGTCCGAGCAACACGACGAGATTC 1980
Qy 1981 AGAGGGCGGACATAAGTGAAGCTTACATCAAGTAGAAGATCAGGACTATGCAAA 2040
Db 1981 AGAGGGCGGACATAAGTGAAGCTTACATCAAGTAGAAGATCAGGACTATGCAAA 2040
Qy 2041 CTCTACTGTATCGCAGAGATTTGATTTCTTCTTTTCTTTTGTCAAATGAAGTCAAAGAT 2100
Db 2041 CTCTACTGTATCGCAGAGATTTGATTTCTTCTTTTCTTTTGTCAAATGAAGTCAAAGAT 2100
Qy 2101 GGGACTCCATGCTCGAGGATAGCCGTAATGTTGTATAGATGGGATATGTGAGAGATT 2160
Db 2101 GGGACTCCATGCTCGAGGATAGCCGTAATGTTGTATAGATGGGATATGTGAGAGATT 2160
Qy 2161 GGATGTGACAAATGCTCTGGATCTGATGCTGTGAAGAGCTGTGGGGTGTGAAGGG 2220
Db 2161 GGATGTGACAAATGCTCTGGATCTGATGCTGTGAAGAGCTGTGGGGTGTGAAGGG 2220
Qy 2221 AATACTCAGCTGCAAGATTCACAGGGGTCTCTACCAAGACCAACCAACCAAG 2280
Db 2221 AATACTCAGCTGCAAGATTCACAGGGGTCTCTACCAAGACCAACCAACCAAG 2280
Qy 2281 TATTATCAGATGGTCAACATTCCTTCTGAGCCCGGAGTATCCGCATCTATGAAGTGAAC 2340
Db 2281 TATTATCAGATGGTCAACATTCCTTCTGAGCCCGGAGTATCCGCATCTATGAAGTGAAC 2340
Qy 2341 GTCTCTACTCTCTAATTTCTGTGGCAATGCCCTCAGAAAGGTACTACCTGAATGGGCAC 2400
Db 2341 GTCTCTACTCTCTAATTTCTGTGGCAATGCCCTCAGAAAGGTACTACCTGAATGGGCAC 2400
Qy 2401 TGGACCGTGGACTGGCCCGGCGGTACAAATTTTCGGGCACTACTTTCGACTACAGACGG 2460
Db 2401 TGGACCGTGGACTGGCCCGGCGGTACAAATTTTCGGGCACTACTTTCGACTACAGACGG 2460

Qy 2461 TCCTATAATGAGCCCGAGAACTTAATCGCTACTGGAACCAACGAGACACTGATTGTG 2520
Db 2461 TCCTATAATGAGCCCGAGAACTTAATCGCTACTGGAACCAACGAGACACTGATTGTG 2520
Qy 2521 GAGCTGCTGTTTTCAGGAAAGGAAACCGGGGTGTTGCTGGGAATACTCCATGCTCGCTTG 2580
Db 2521 GAGCTGCTGTTTTCAGGAAAGGAAACCGGGGTGTTGCTGGGAATACTCCATGCTCGCTTG 2580
Qy 2581 GGGACCGAGAGCAGCCCTCGCCAGCCAGCTACATTTGGGCCCATCGTGGCTCTGAG 2640
Db 2581 GGGACCGAGAGCAGCCCTCGCCAGCCAGCTACATTTGGGCCCATCGTGGCTCTGAG 2640
Qy 2641 TGTCTCGTGTCTGCGGAGGGGACAGATGACCGTGTGAGAGAGGGCTGCTCAGAGACCTG 2700
Db 2641 TGTCTCGTGTCTGCGGAGGGGACAGATGACCGTGTGAGAGAGGGCTGCTCAGAGACCTG 2700
Qy 2701 AAGTTTCAAGTAATAATATCTTCTGCAATCCCAAGACAAGACCTGTCAACGGGCTGTG 2760
Db 2701 AAGTTTCAAGTAATAATATCTTCTGCAATCCCAAGACAAGACCTGTCAACGGGCTGTG 2760
Qy 2761 CCTTGCAGAAATATCTGCTGTCTCCAGCTGTGCTCCGTTGGGAACTGAGTGTGCTGAGT 2820
Db 2761 CCTTGCAGAAATATCTGCTGTCTCCAGCTGTGCTCCGTTGGGAACTGAGTGTGCTGAGT 2820
Qy 2821 CGGACGTGTGCGGGGGTGGCCAGAGCGGCCCTGTGCTGCAATCCCAAGACAAGACCTGTCA 2880
Db 2821 CGGACGTGTGCGGGGGTGGCCAGAGCGGCCCTGTGCTGCAATCCCAAGACAAGACCTGTCA 2880
Qy 2881 GACTCGAGACAGTTCCTGGGCGAGCTGTGCTGCTCAGCTGTCTCCCTCAGCTGTCTCCCTC 2940
Db 2881 GACTCGAGACAGTTCCTGGGCGAGCTGTGCTGCTCAGCTGTCTCCCTCAGCTGTCTCCCTC 2940
Qy 2941 TGCAACTCTCAGAGCTGTCCCACTGTGAGAGCGGCCCTGTGGGCAAGTGTCTCACAC 3000
Db 2941 TGCAACTCTCAGAGCTGTCCCACTGTGAGAGCGGCCCTGTGGGCAAGTGTCTCACAC 3000
Qy 3001 ACCTGTGGGAAGGGGTGAGGAGAGCGGGCAGTGGCTGTAAAGACCAACCCCTCGGCC 3060
Db 3001 ACCTGTGGGAAGGGGTGAGGAGAGCGGGCAGTGGCTGTAAAGACCAACCCCTCGGCC 3060
Qy 3061 AGAGCGCAGCTGTGCTGCGCAGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120
Db 3061 AGAGCGCAGCTGTGCTGCGCAGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120
Qy 3121 GCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3180
Db 3121 GCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3180
Qy 3181 TCCCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3240
Db 3181 TCCCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3240
Qy 3241 AAGTATGTTTCTGGAAGTATCGAGAGCTGCGCTCAAGAAAGTGTCTTAAATGTGCTGAA 3300
Db 3241 AAGTATGTTTCTGGAAGTATCGAGAGCTGCGCTCAAGAAAGTGTCTTAAATGTGCTGAA 3300
Qy 3301 CCCAGCTGTGAGCTGGAACGCTGCGCCCGCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3360
Db 3301 CCCAGCTGTGAGCTGGAACGCTGCGCCCGCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3360
Qy 3361 GCTGCTGCGGAGCCCTCGAGGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420
Db 3361 GCTGCTGCGGAGCCCTCGAGGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420
Qy 3421 AGCTGTGGGGGAGCGTTCAGAGAGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3480
Db 3421 AGCTGTGGGGGAGCGTTCAGAGAGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3480
Qy 3481 TCAGGCTGCTCTGCTGCAACAGAGAGCTTCGCGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 3540
Db 3481 TCAGGCTGCTCTGCTGCAACAGAGAGCTTCGCGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 3540

QY 3541 CCCATTGCAGAGAGAAAGTGGCTTCTGCAAGAGACTACTTCCACTGGTGCTACCTGGTA 3600
Db 3541 CCCATTGCAGAGAGAAAGTGGCTTCTGCAAGAGACTACTTCCACTGGTGCTACCTGGTA 3600
QY 3601 CCCAGCAGGGATGTGCGAGCCCAAGATTCTACGGCAAGCAGTGTGCGAAGACTTGTCTCT 3660
Db 3601 CCCAGCAGGGATGTGCGAGCCCAAGATTCTACGGCAAGCAGTGTGCGAAGACTTGTCTCT 3660
QY 3661 AGTCCAACTTGTGA 3675
Db 3661 AGTCCAACTTGTGA 3675

RESULT 8

US-10-399-645-19
; Sequence 19, Application US/10399645
; Publication No. US20040029249A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; LEE, Ernestine A.
; APPLICANT: HAFALIA, April J.A.; YOE, Henry
; APPLICANT: LAL, Preeti G.; YAO, Monique G.
; APPLICANT: LU, Yan; CHAWLA, Narinder K.
; APPLICANT: WARREN, Bridget A.; LU, Dyung Aina M.
; APPLICANT: BAUGHN, Mariah R.; DELEGEANE, Angelo M.
; APPLICANT: BURFORD, Neil; BOWSKI, Mark L.
; APPLICANT: LEE, Sally; XU, Yuming
; APPLICANT: GRIFFIN, Jennifer A.; KALLICK, Deborah A.
; APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.
; APPLICANT: ISON, Craig H.; TANG, Y. Tom
; APPLICANT: AZIMZAI, Yalda; ELLIOTT, Vicki S.
; APPLICANT: SWARNAKAR, Anita; RAMKOMAR, Jayaixmi
; APPLICANT: NGUYEN, Daniel B.; TRIBOULEY, Catherine M.
; APPLICANT: LO, Terence P.; AU-YOUNG, Janice K.
; APPLICANT: THANGAVELU, Kavitha; KEARNEY, Liam
; TITLE OF INVENTION: PROTEASES
; FILE REFERENCE: PI-0263 USN
; CURRENT APPLICATION NUMBER: US/10/399,645
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: PCT/US01/51034
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: US 60/241,573
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: US 60/243,643
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: US 60/245,256
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/248,395
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: US 60/249,826
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252,303
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,981
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PERL Program
; SEQ ID NO 19
; LENGTH: 4888
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040029249A1 3533147CB1
US-10-399-645-19

Query Match 85.3%; Score 3135.2; DB 17; Length 4888;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3163; Conservative 0; Mismatches 8; Indels 3; Gaps 2;
QY 502 TCAGGCATGATACGAACAGAGAGGCGAGATTACTTCTTAAGGCCACTTCTTACACACCTC 561
Db 541 TCAGGCATGATACGAACAGAGAGGCGAGATTACTTCTTAAGGCCACTTCTTACACACCTC 600

QY 562 TCATGGAACCTCGGAGAGAGCTGCCAAGGCGAGCTGCCCATCCACGCTACTGTACAAGAGA 621
Db 601 TCATGGAACCTCGGAGAGAGCTGCCAAGGCGAGCTGCCCATCCACGCTACTGTACAAGAGA 660
QY 622 TCACAGAGCCCCATGCTCTCTGGGGCCAGTGAAGTCTCTGGTGAACCTCAAGAGCATGGGAG 681
Db 661 TCACAGAGCCCCATGCTCTCTGGGGCCAGTGAAGTCTCTGGTGAACCTCAAGAGCATGGGAG 720
QY 682 CTGGCACAATCAACCCCTGCAAGAGAGAGCTTTCGCTGGGAGCTGCCACAAAAGCAGCAT 741
Db 721 CTGGCACAATCAACCCCTGCAAGAGAGAGCTTTCGCTGGGAGCTGCCACAAAAGCAGCAT 780
QY 742 TTCTGTGGAAGAGCGCAAGAAATACATGCCCGGAGCTGCCAAGAGAGAGCTTTCATCTT 801
Db 781 TTCTGTGGAAGAGCGCAAGAAATACATGCCCGGAGCTGCCAAGAGAGAGCTTTCATCTT 840
QY 802 CCAGATGAGTATAAGTCTTTCGCTTACGGCATAGAGCGCTCTCTCTGAGGTCCTCATAGAAAT 861
Db 841 CCAGATGAGTATAAGTCTTTCGCTTACGGCATAGAGCGCTCTCTCTGAGGTCCTCATAGAAAT 900
QY 862 GAAGAACTGAACCTGAGAGACCTTGTGTGGTTCGACAAAAGAGATGACAAAACCATGGC 921
Db 901 GAAGAACTGAACCTGAGAGACCTTGTGTGGTTCGACAAAAGAGATGACAAAACCATGGC 960
QY 922 CATGAAAATATCACCACTACGCTCAGCATACTCAACATGGTATCTGCTTTATTCAAA 981
Db 961 CATGAAAATATCACCACTACGCTCAGCATACTCAACATGGTATCTGCTTTATTCAAA 1020
QY 982 GATGGAACAATAGGAGGAACATCAACATTTGCAATTTGAGGTCGTAGTCTCTCTAGAGAT 1041
Db 1021 GATGGAACAATAGGAGGAACATCAACATTTGCAATTTGAGGTCGTAGTCTCTCTAGAGAT 1080
QY 1042 GAACAGCCAGAGCTGGTGATAAGTCAACAGCAGAGACCAACCTTTAAGTAGTCTTCTGCCAG 1101
Db 1081 GAACAGCCAGAGCTGGTGATAAGTCAACAGCAGAGACCAACCTTTAAGTAGTCTTCTGCCAG 1140
QY 1102 TGGAGTCTGGATTTGATGGGAAAAGATGGGACTCTGTCATGACCAACGCCATCTTACTGACT 1161
Db 1141 TGGAGTCTGGATTTGATGGGAAAAGATGGGACTCTGTCATGACCAACGCCATCTTACTGACT 1200
QY 1162 GGTCTGGATATATGTTCTCTGGAAGATGAGCCCTGTGACACTTTGGGATTTGACCCATA 1221
Db 1201 GGTCTGGATATATGTTCTCTGGAAGATGAGCCCTGTGACACTTTGGGATTTGACCCATA 1260
QY 1222 AGTGGAAATGTTAGTAAATATCGCAGCTGCAGATTAATGAAGATACAGGTCTTGGAGCTG 1281
Db 1261 AGTGGAAATGTTAGTAAATATCGCAGCTGCAGATTAATGAAGATACAGGTCTTGGAGCTG 1320
QY 1282 GCCTTCACCATTTGCCCATGAGTCTGGACACAACTTTTGGCATGATTCATGATGGAGAGGG 1341
Db 1321 GCCTTCACCATTTGCCCATGAGTCTGGACACAACTTTTGGCATGATTCATGATGGAGAGGG 1380
QY 1342 AACATGTTGTAAGAGTCCGAGGCGACATCATGTCCCTACATTTGGCAGAGCGCAATGGA 1401
Db 1381 AACATGTTGTAAGAGTCCGAGGCGACATCATGTCCCTACATTTGGCAGAGCGCAATGGA 1440
QY 1402 GTCTTCTCTGTGCTCAACCTGCGAGCGCCAGATCTACACAAAATTTCTAAGACCGCTCAA 1461
Db 1441 GTCTTCTCTGTGCTCAACCTGCGAGCGCCAGATCTACACAAAATTTCTAAGACCGCTCAA 1500
QY 1462 GCTATCTGCTTGTGATCAGCCAAAGCCTGTGAAGAGATAAAGATCTCTGAGAAATTTG 1521
Db 1501 GCTATCTGCTTGTGATCAGCCAAAGCCTGTGAAGAGATAAAGATCTCTGAGAAATTTG 1560
QY 1522 CCAGGAGAAATATGATGCAACACACAGTCCAGTGGAGTTCGAGAGAGAGCGCAAG 1581
Db 1561 CCAGGAGAAATATGATGCAACACACAGTCCAGTGGAGTTCGAGAGAGAGCGCAAG 1620
QY 1582 CTCTGATGCTGCACTTTTAAAAAGGACATCTCTAAAGCCCTGTGGTGCCATCGTATTGGA 1641
Db 1621 CTCTGATGCTGCACTTTTAAAAAGGAGCATCTGTAAAGCCCTGTGGTGCCATCGTATTGGA 1680
QY 1642 AGGAAATGTGAGACTAAATTTTATGCCAGCAGAGAGGACAAATTTGTGGGCATGACATG 1701

Db 1681 AGGAAATGTGAGACTAAATTTATGCCAGAGCAGAGAGGACAAATTTGTGGCATGACATG 1740
Qy 1702 TGGTCCGGGAGGAGCAGTGTGTGAATAATGGTATGAAGGCCCCCAAGCCACCCCATGGC 1761
Db 1741 TGGTCCGGGAGGAGCAGTGTGTGAATAATGGTATGAAGGCCCCCAAGCCACCCCATGGC 1800
Qy 1762 CACTGTGCGAGTGGTCTTCTGGTCCCATGCTCCAGGACCTGCGGAGGGGAGTATCT 1821
Db 1801 CACTGTGCGAGTGGTCTTCTGGTCCCATGCTCCAGGACCTGCGGAGGGGAGTATCT 1860
Qy 1822 CATAGGAGTCCCTCTGCAACCAAGCCCATCGCATGGAGGGAAGTTCTGTGAGGGC 1881
Db 1861 CATAGGAGTCCCTCTGCAACCAAGCCCATCGCATGGAGGGAAGTTCTGTGAGGGC 1920
Qy 1882 TCCACTCGCACTCTGAAGCTCTGCACAGTCAAGAAATGTCCCGGAGCAGTGTGACTTC 1941
Db 1921 TCCACTCGCACTCTGAAGCTCTGCACAGTCAAGAAATGTCCCGGAGCAGTGTGACTTC 1980
Qy 1942 CGTGTCTCAGTGTGCCGAGCAACAGCAGACGATTCAGAGGGCGGCACTACAGTGG 2001
Db 1981 CGTGTCTCAGTGTGCCGAGCAACAGCAGACGATTCAGAGGGCGGCACTACAGTGG 2040
Qy 2002 AAGCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAACTCTACTGTATCGCAGAAGGA 2061
Db 2041 AAGCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAACTCTACTGTATCGCAGAAGGA 2100
Qy 2062 TTGTGATTTCTCTTTTGTTCCTTGTCAAAATGAAGTCAAGATGGGACTCCATGCTCGAGGAT 2121
Db 2101 TTGTGATTTCTCTTTTGTTCCTTGTCAAAATGAAGTCAAGATGGGACTCCATGCTCGAGGAT 2160
Qy 2122 AGCGTAATGTTGTATAGATGGATATGTGAGAGATTTGGATGTGACAAATGTCCTTGA 2181
Db 2161 AGCGTAATGTTGTATAGATGGATATGTGAGAGATTTGGATGTGACAAATGTCCTTGA 2220
Qy 2182 TCTGATCTGTGTAAGACGTCTGTGGGGTGTGTAACGGGAATAACTCAGCCTGCACGATT 2241
Db 2221 TCTGATCTGTGTAAGACGTCTGTGGGGTGTGTAACGGGAATAACTCAGCCTGCACGATT 2280
Qy 2242 CACAGGGTCTCTACACCAAGCACCACACCCACCACTATATACAGTGTGACCATTT 2301
Db 2281 CACAGGGTCTCTACACCAAGCACCACCACTATATACAGTGTGACCATTT 2340
Qy 2302 CTTTGTGAGCCGGAGTATCCGATCTATGAATGAACGTCTCTACCTCTACATTTCT 2361
Db 2341 CTTTGTGAGCCGGAGTATCCGATCTATGAATGAACGTCTCTACCTCTACATTTCT 2400
Qy 2362 GTGCGCAATGCCCTCAGAGGTACTACTGTAATGGGCACTGACCGTGAATGGCCCGGC 2421
Db 2401 GTGCGCAATGCCCTCAGAGGTACTACTGTAATGGGCACTGACCGTGAATGGCCCGGC 2460
Qy 2422 CGGTACAAATTTTTCGGGCACTACTTTTCGACTACAGACGGTCTCTAATAGAGCCGAGAAC 2481
Db 2461 CGGTACAAATTTTTCGGGCACTACTTTTCGACTACAGACGGTCTCTAATAGAGCCGAGAAC 2520
Qy 2482 TTAAATCGTACTGGAACCAACCAAGAGACACTGATTTGGAGTCTGTTTCAGGGGAGG 2541
Db 2521 TTAAATCGTACTGGAACCAACCAAGAGACACTGATTTGGAGTCTGTTTCAGGGGAGG 2580
Qy 2542 AACCCGGGTGTCCTGGGAATCTCCATGCTCGCTTTGGGACCGAGAGCAGCCCT 2601
Db 2581 AACCCGGGTGTCCTGGGAATCTCCATGCTCGCTTTGGGACCGAGAGCAGCCCT 2640
Qy 2602 GCCCAGCCAGCTACACTTTGGGCCATCGTGGCTCTGAGTGTCTCGTGTCTCTGGAGGG 2661
Db 2641 GCCCAGCCAGCTACACTTTGGGCCATCGTGGCTCTGAGTGTCTCGTGTCTCTGGAGGG 2700
Qy 2662 GCACAGATGACCGTGAAGAGGGCTGTACAGAGACCTGAAGTTTCAAGTAAATATGTCC 2721
Db 2701 GCACAGATGACCGTGAAGAGGGCTGTACAGAGACCTGAAGTTTCAAGTAAATATGTCC 2760
Qy 2722 TTCTGCAATCCCAAGACGACTCTCAGCGGGCTGTGCTGCAAGATATCTGCTGT 2781

Db 2761 TTCTGCAATCCCAAGACACGACCTGTACGGGGCTGGTGCCTTTCGAAAGTATCTGCTGT 2820
Qy 2782 CTTCCAGCTGTCTCGTGGGAACTGGAGTGTCTGAGTCGACGTCGAGTGTGGCGGGTGGC 2841
Db 2821 CTTCCAGCTGTCTCGTGGGAACTGGAGTGTCTGAGTCGACGTCGAGTGTGGCGGGTGGC 2880
Qy 2842 CAGAGCGCCCGTGTGAGTGTGACACAGGGGGTGTCTATGACTCGAGACGAGTCCCGGC 2901
Db 2881 CAGAGCGCCCGTGTGAGTGTGACACAGGGGGTGTCTATGACTCGAGACGAGTCCCGGC 2940
Qy 2902 AGCCTGTGCTCAGCTGTCTCCAGCAGGAGGCTGTGCAACTCTCAGAGTGTCCCA 2961
Db 2941 GGCTGTGCTCTAG-CTGGTCCCTCCAGCAGGAGGCTGTCACTCTCAGAGTGTCCCA 2999
Qy 2962 CTTGATGAGAGCGCCGGCCCTGGGCGAGAGTGTCTCACACACTGTGTGGGAAGGGTGGAGG 3021
Db 3000 CTTGATGAGAGCGCCGGCCCTGGGCGAGAGTGTCTCACACACTGTGTGGGAAGGG--TGGAG 3057
Qy 3022 AAGCGGCGAGTGGCTGTGAAGACACCAACCCCTCGGCCAGAGCGAGTGTGTGCCCGAC 3081
Db 3058 AAGCGGCGAGTGGCTGTGAAGACACCAACCCCTCGGCCAGAGCGAGTGTGTGCCCGAC 3117
Qy 3082 GCTGTCTGCACCTCGAGCCCAAGCCAGGATGCAATGAAGCTGTCTCAGCGCTGC 3141
Db 3118 GCTGTCTGCACCTCGAGCCCAAGCCAGGATGCAATGAAGCTGTCTCAGCGCTGC 3177
Qy 3142 CACAAGCCAAAGAGCTCAGTGTGTGTCCGCTGGTCCAGTGTCTGTGACATGT 3201
Db 3178 CACAAGCCAAAGAGCTCAGTGTGTGTCCGCTGGTCCAGTGTCTGTGACATGT 3237
Qy 3202 GAAAGAGAAACACAGAAAGATTTCTTAAATGTGCTGAAAAGATGTTTCTGGAAGTAT 3261
Db 3238 GAAAGAGAAACACAGAAAGATTTCTTAAATGTGCTGAAAAGATGTTTCTGGAAGTAT 3297
Qy 3262 CGAGAGCTGGCTCAAGAGAGTGTCTCATTTGCGAAGCCAGCTGTGAGCTGGAAGCT 3321
Db 3298 CGAGAGCTGGCTCAAGAGAGTGTCTCATTTGCGAAGCCAGCTGTGGAAGCT 3357
Qy 3322 GCCTGGCCCGCTTCATGCCCCAGGCACCCCCATTTGTGTGTGGGAGCCCTCGAGG 3381
Db 3358 GCCTGGCCCGCTTCCATGCCCGAGGCACCCCCATTTGTGTGTGGGAGCCCTCGAGG 3417
Qy 3382 GCAGAGCTGTTGTCTCAACCTGTCTCAGTGCAGCGCAGCTGTGTGGGGAGGGTTCAG 3441
Db 3418 GCAGAGCTGTTGTCTCAACCTGTCTCAGTGCAGCGCAGCTGTGTGGGGAGGGTTCAG 3477
Qy 3442 ACAGAGTCCGTGAGTGTGCTGGTGGGGCGCGCGCTCAGAGTGTCTCCTGACCCAG 3501
Db 3478 ACAGAGTCCGTGAGTGTGCTGGTGGGGCGCGCGCTCAGAGTGTCTCCTGACCCAG 3537
Qy 3502 AAGCCTTTCGGCTCCCTGGCTGCAACACTCCTCTTCTGCCCCATTGCAAGAAAGAT 3561
Db 3538 AAGCCTTTCGGCTCCCTGGCTGCAACACTCCTCTTCTGCCCCATTGCAAGAAAGAT 3597
Qy 3562 GCCTTCTGCAAGACTACTTCCACTGTGTGTCTGCTGTACCCAGCAGCGGATGTGAGC 3621
Db 3598 GCCTTCTGCAAGACTACTTCCACTGTGTGTCTGCTGTACCCAGCAGCGGATGTGAGC 3657
Qy 3622 CACAAGTCTTACGCAAGCAGTGTGTGCAAGACTTGTCTTAAGTCCAACTGTGA 3675
Db 3658 CACAAGTCTTACGCAAGCAGTGTGTGCAAGACTTGTCTTAAGTCCAACTGTGA 3711

RESULT 9

US-09-981-151A-1

; Sequence 1, Application US/09981151A

; Publication No. US20030212256A1

; GENERAL INFORMATION:

; APPLICANT: Edinger, Shlomit R

; APPLICANT: Gerlach, Valerie

; APPLICANT: MacDougall, John R

; APPLICANT: Malyankar, Muriel M

; APPLICANT: Smithson, Glenda

```
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A
; APPLICANT: Stone, David J
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Shimkets, Richard A
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Gangolli, Esha R
; APPLICANT: Fernandes, Elma R
; APPLICANT: Gorman, Linda
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-168
; CURRENT APPLICATION NUMBER: US/09/981,151A
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/241,040
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,058
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,063
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,243
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/242,152
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/242,482
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,611
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,612
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,880
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,881
; PRIOR FILING DATE: 2000-10-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2997
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (857)..(858)
; OTHER INFORMATION: Wherein n is an a or t or c or g.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2383)
; OTHER INFORMATION: Wherein n is an a or t or c or g.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2983)
; OTHER INFORMATION: wherein n is an a or t or c or g.
; US-09-981-151A-1

Query Match 55.9%; Score 2055.6; DB 10; Length 2997;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 2386; Conservative 0; Mismatches 182; Indels 143; Gaps 8;

QY 69 GCAGGCACCTGCGGCCATGGGACCCGCGAGCGGCGGCTGGGAGCCCGGAGCGTCCC 128
Db 202 GCAGGCACCTGCGGCCATGGGACCCGCGAGCGGCGGCTGGGAGCCCGGAGCGTCCC 261
QY 129 GCGTCTCTTCCACCCGCGAGCGGCGGCTGGATGGAAGGGCG----- 175
Db 262 GCGTCTCTTCCACCCGCGAGCGGCGGCTGGATGGAAGGGCG----- 175

; 176 ----- 175
322 GCTGGAACCATCGTTCTCAGCAAACTAAACACAGGAACAGAAAAACAAACACATGCGATGTT 381
176 -----AATATGACCTGGTCTCTGCTACGAGGTTGACACAGGGGCGGATTAAGTGTCC 228
382 CTCACCTCAATATAGACCTGGTCTCTGCTACGAGGTTGACCAAGGGGCGGATTAAGTGTCC 441
229 CATGAATCATGCACCATCAGCGGCGGAGAGAGCAGTGGCCGTGTCCGAGGTTGAGTCT 288
442 CATGAATCATGCACCATCAGCGGCGGAGAGAGCAGTGGCCGTGTCCGAGGTTGAGTCT 501
289 CTTACCTTCGGCTGAAAGGCTCCAGGCAAGCTTCCAGTGTGATCTGAGGACTTCCAGC 348
502 CTTACCTTCGGCTGAAAGGCCCCAGGCAAGCTTCCAGTGTGATCTGAGGACTTCCAGC 561
349 AGCCTAGTGGCTCTGGCTTTATTGTGACAGAGTGTGGGAAAGACAGGCACTAAGTCTGTG 408
562 AGCCTAGTGGCTCTGGCTTTATTGTGACAGAGTGTGGGAAAGACAGGCACTAAGTCTGTG 621
409 CAGACTTTACCGCCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGCATCACACAGAAAC 468
622 CAGACTTTACCGCCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGCATCACACAGAAAC 681
469 TCCTCA-----GTGGCCCTTTCAACCTGCCAAGGCTGTGCA 504
682 TCGCCATCGCATGGAGGGAAGTTCTGTGAGGGCTCCCACTCGCACTCTGAAGCTCTGCAC 741
505 GGCATGATACGAAACAGAGAGGCGAGATTACTTCTTAAGGCCACTTCTTTCACACCTCTCA 564
742 AGTCAGAAATGTCCCCGGGACAGTGTGACTTCTTCGCTGCTCAGTGTGCGGAGCAAC 801
565 TGGAAACTCGGAGAGCTGCCAAGGCGAGTGGCCATCCACGCTACTGTACAGAGAGATCC 624
802 AGCAGACGATTGAGAGGGCGGCACTACAAGTGG--AAGCTTTACACTCAAGTGAAGNNG 859
625 ACAGAGCCCCATGCTCTCGGGGCCAGTGAAGTCTCTGGTGACCTC-----AAGG 672
860 ACTTATGCAAACTCTACTGTATGCGAGAAGGATTTGATTTCTTTCTTTGTCAATA 919
673 ACATGGGAGCTGGCAGCATCAACCCCTGCGACAGAGGAGCTTCGCTGGGAGCTGCCACAA 732
920 AAGTCAAGATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTGTATAGATGGGATAT 979
733 AAGCAGCATTTCTGTGGAAGAGCGAAGAAATACATGCCCGCAGCTCCCAAGGAAGACCTC 792
980 GTGAGCTCAGTGTGGTGTCCACATCTG-CGCACATGCCCGCAGCTCCCAAGGAAGACCTC 1038
793 TTCATCTTCCAGATGAGTATAAGTCTTTCAGGCATAAGCGCTCTCTTCTGAGGTCC 852
1039 TTCATCTTCCAGATGAGTATAAGTCTTTCAGGCATAAGCGCTCTCTTCTGAGGTCC 1098
853 CATAGAAATGAAGAACTGAACGTGGAGACTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 912
1099 CATAGAAATGAAGAACTGAACGTGGAGACTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1158
913 AACCATGGCCATGAAATATCACCACTACGCTGCTCACGATCTCAACATGGTATCTGCT 972
1159 AACCATGGCCATGAAATATCACCACTACGCTGCTCACGATCTCAACATGGTATCTGCT 1218
973 TTATTCAAAGATGGAAACAATAGGAGGAAACATCAACATTTGCAATTTAGTGTCTGATTTCT 1032
1219 TTATTCAAAGATGGAAACAATAGGAGGAAACATCAACATTTGCAATTTAGTGTCTGATTTCT 1278
1033 CTAGAAAGATGAACAGCCAGGACTGTGTGATAAGTCAACAGCAGACACACACCTTAAGTAGC 1092
1279 CTAGAAAGATGAACAGCCAGGACTGTGTGATAAGTCAACAGCAGACACACACCTTAAGTAGC 1338
1093 TTCTGCGAGTGGCGTCTCGATTGTGTTGGGGAAGATGGGAGCTGCTCATGACCAAGCCATC 1152
1339 TTCTGCGAGTGGCGTCTCGATTGTGTTGGGGAAGATGGGAGCTGCTCATGACCAAGCCATC 1398
```


; PRIOR APPLICATION NUMBER: 60/242,482
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,611
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,612
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,880
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,881
; PRIOR FILING DATE: 2000-10-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2902
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-981-151A-5

Query Match 55.0%; Score 2022.8; DB 10; Length 2902;
Best Local Similarity 89.0%; Pred. No. 0;
Matches 2323; Conservative 0; Mismatches 167; Indels 120; Gaps 8;

QY 69 GCAGGACCTGCGTGGCCATGGACCCGCGAGCGGCGCTGGAGCCGCGAGCGTCCC 128
Db |||||
QY 163 GCAGGACCTGCGTGGCCATGGGACCCGCGAGCGGCGCTGGAGCCGCGAGCGTCCC 222
Db |||||
QY 129 GCGTCTCTCTCCACCGCGAGCGGCGGCGCTGGATGGAAAGGCGGAATATGACCTGGT 188
Db |||||
QY 223 GCGTCTCTCTCCA-----GACCTGGT 243
Db |||||
QY 189 CTCTGCTTACGAGTTGACACAGGGGCGATTAAGTGTCCCATGAAATCATGCAACATCA 248
Db |||||
QY 244 CTCTGCTTACGAGTTGACACAGGGGCGATTAAGTGTCCCATGAAATCATGCAACATCA 303
Db |||||
QY 249 GCGGCGGAGAGAGAGTGGCGTGGCGAGTTGAGTCTCTTCACTTCCGCTGGAAGG 308
Db |||||
QY 304 GCGGCGGAGAGAGAGTGGCGTGGCGAGTTGAGTCTCTTCCAGGTATGCGAG 363
Db |||||
QY 309 CTCACGGCAGACTTCCAGCTGGATCTGAGGACTTCCA-----GCAGCTAGTGG 358
Db |||||
QY 364 AGCCAGAGAGCTCAGACTGTGTGGAGGCGCTTCCCATGCTAAATCTCAGCGGGGT 423
Db |||||
QY 359 CTCTGGCTTTATGTGCGAGAGTGGGAAAGACAGGCACTAAGTCTGTGCGAGCTTTAC 418
Db |||||
QY 424 TTGGAACCTTTCCAAATGTTGCTCTCACTGGAGGGAACAGCATGCTTCCAAGAGATAAT 483
Db |||||
QY 419 CGCCAGAGAGACTTCTGTCTATCAAGGCTTTTGGGATCACACAG--AACTCTCAGT 476
Db |||||
QY 484 AACAAATGCAATGCTTGGAGAAATCGGCCCTGGCTTCAACCGAGAAAGTCTAAATTTGTGTTTT 543
Db |||||
QY 477 GGCCTTTTCAACTGCGCAAGGCTTGTGAGGATGATACGAAACAGAGAGGCGAGATTACTT 536
Db |||||
QY 544 CTTTCTTTCTTTATTTTTCAGTCAAGGATGATACGAAACAGAGAGGCGAGATTACTT 603
Db |||||
QY 537 CCTAAGGCGCACTTCTTCAACCTCTCATGAAACTCGGAGAGCTGCCCAAGGCGAGCTC 596
Db |||||
QY 604 CCTAAGGCGCACTTCTTCAACCTCTCATGAAACTCGGAGAGCTGCCCAAGGCGAGCTC 663
Db |||||
QY 597 GGCATCCAGTACTGTACAGAGATCCACAGAGCCCATGCTCTCGGGGCCAGTGAAGT 656
Db |||||
QY 664 GGCATCCAGTACTGTACAGAGATCCACAGAGCCCATGCTCTCGGGGCCAGTGAAGT 693
Db |||||
QY 657 CTGTGTGACTCAAGGACATGGAGCTGGCAGATCAACCCCTGCAGAGGCGAGCTTCG 716
Db |||||
QY 694 CTTGTGTACTCAAGGACATGGAGCTGGGAGCTGGGACATCAACCCCTGCACAGAGGAGCTTCG 753
Db |||||
QY 717 CTGTGGACTGCCAAGAGAGAGATTTCTGTGGAAGACGCAAGAAATACATGCCCGAGCC 776
Db |||||
QY 754 CTTGGGACTGCCAAGAGAGAGATTTCTGTGGAAGACGCAAGAAATACATGCCCGAGCC 813
Db |||||
QY 777 TCCAAAGGAGACTCTTCACTTGGCCAGATGATTAAGTCTTGTGCTTACGGCATGAAGG 836
Db |||||

Db 814 TCCCAAGGAAGACCTCTTTCATCTTGGCCAGATGAGTATAAGTCTTGTCTTACGGCATGAAGCG 873
QY 837 CTCTCTTCTGAGGTCCCATAGAAATGAAGAACTGAACGTTGAGACCTTGGTGTGTCGA 896
Db |||||
QY 874 CTCTCTTCTGAGGTCCCATAGAAATGAAGAACTGAACGTTGAGACCTTGGTGTGTCGA 933
Db |||||
QY 897 CAARAAGATGATGCAAAAGCATGGCCATGAAAATATCACCACCTAGCTGTCTCAGGATCT 956
Db |||||
QY 934 CAARAAGATGATGCAAAAGCATGGCCATGAAAATATCACCACCTAGCTGTCTCAGGATCT 993
Db |||||
QY 957 CAACATGGTATCTGCTTTTATTCAAAGATGGAACAAATAGGAGGAAACATCAACATTTGCAAT 1016
Db |||||
QY 994 CAACATGGTATCTGCTTTTATTCAAAGATGGAACAAATAGGAGGAAACATCAACATTTGCAAT 1053
Db |||||
QY 1017 TGTAGGTCTGATTTCTTTAGAAAGATGAACAGCCAGCACTGGTGTATAGTCAACACGAGA 1076
Db |||||
QY 1054 TGTAGGTCTGATTTCTTTAGAAAGATGAACAGCCAGGACTGGTGTATAGTCAACACGAGA 1113
Db |||||
QY 1077 CCACACCTTAAGTAGTCTTCTGCCAGTGGCAGTCTGAAATGATGGGAAAGATGGGACTCG 1136
Db |||||
QY 1114 CCACACCTTAAGTAGTCTTCTGCCAGTGGCAGTCTGAAATGATGGGAAAGATGGGACTCG 1173
Db |||||
QY 1137 TCATGACCAACGCGCATCTTACTGACTGCTCTGGATATATGTTCTCGAAGAAATCAGCCCTG 1196
Db |||||
QY 1174 TCATGACCAACGCGCATCTTACTGACTGCTCTGGATATATGTTCTCGAAGAAATCAGCCCTG 1233
Db |||||
QY 1197 TGACATTTTGGGATTTGCAACCATTAAGTGGAAATGTTAGTAAATATCGCAGCTGCAAGAT 1256
Db |||||
QY 1234 TGACATTTTGGGATTTGCAACCATTAAGTGGAAATGTTAGTAAATATCGCAGCTGCAAGAT 1293
Db |||||
QY 1257 TAATGAGATACAGGTCTTGGACTGGCCTTCCACCATTGGCCATGAGTCTGGACACAACTT 1316
Db |||||
QY 1294 TAATGAGATACAGGTCTTGGACTGGCCTTCCACCATTGGCCATGAGTCTGGACACAACTT 1353
Db |||||
QY 1317 TGSCATGATTTCAATGATGAGAGAGGGAACATGTTGTAAGAAAGTCCGAGGCAACATCATGTC 1376
Db |||||
QY 1354 TGSCATGATTTCAATGATGAGAGAGGGAACATGTTGTAAGAAAGTCCGAGGCAACATCATGTC 1413
Db |||||
QY 1377 CCCTACATTTGGCAGGACGCAATGGAGTCTTCTCTGGTCACTCGCAGCGCCAGATCT 1436
Db |||||
QY 1414 CCCTACATTTGGCAGGACGCAATGGAGTCTTCTCTGGTCACTCGCAGCGCCAGATCT 1473
Db |||||
QY 1437 ACACAAATTTCTAAGCACCGCTCAAGCTATCTGCTTCTGCTGATCAGCCAAAGCCTGTGA 1496
Db |||||
QY 1474 ACACAAATTTCTAAGCACCGCTCAAGCTATCTGCTTCTGCTGATCAGCCAAAGCCTGTGA 1533
Db |||||
QY 1497 GGAATACAGATGATCTTGAGAAATTTCCAGAGAAATTTATATGATGCAAAACACACAGTGC 1556
Db |||||
QY 1534 GGAATACAGATGATCTTGAGAAATTTCCAGAGAAATTTATATGATGCAAAACACACAGTGC 1593
Db |||||
QY 1557 GTGGCAGTTCCGAGAGAAAGCAAGCTCTGCTGCTGGACTTTTAAAGAGGACATCTGTAA 1616
Db |||||
QY 1594 GTGGCAGTTCCGAGAGAAAGCAAGCTCTGCTGCTGGACTTTTAAAGAGGACATCTGTAA 1653
Db |||||
QY 1617 AGCCCTGTGGTGCATCTGTAATTTGGAAGAAATTTGAGACTAAATTTATGCGCAGCAGAGA 1676
Db |||||
QY 1654 AGCCCTGTGGTGCATCTGTAATTTGGAAGAAATTTGAGACTAAATTTATGCGCAGCAGAGA 1713
Db |||||
QY 1677 AGGCACAATTTCTGCGCATGACATGTTGGTGGCGGAGGACAGTGTGAAATATGCTGA 1736
Db |||||
QY 1714 AGGCACAATTTCTGCGCATGACATGTTGGTGGCGGAGGACAGTGTGAAATATGCTGA 1770
Db |||||
QY 1737 TGAAGGCCCCAAGCCACACCCATGGCCACTGGTGGAGCTGGTCTTCTTGTGTCGCCATGCTC 1796
Db |||||
QY 1771 TGAAGGCCCCAAGCCACACCCATGGCCACTGGTGGAGCTGGTCTTCTTGTGTCGCCATGCTC 1830
Db |||||
QY 1797 CAGGACCTCGCGAGGAGGAGTATCTCATAGGAGTGC-----CTCTGCACCAACCCCAAGCC 1853
Db |||||
QY 1831 CAGGACCTCGCGAGGAGGAGTATCTCATAGGAGTGCCTCTCAAAATATACATTTCCAGGCC 1890
Db |||||
QY 1854 ATCGCATGAGAGGAGGAGTCTGTGAGGCTCCACTGCACTCTGAAGCTCTGCAACAGTCA 1913
Db |||||
QY 1891 ATCGCATGAGAGGAGGAGTCTGTGAGGCTCCACTGCACTCTGAAGCTCTGCAACAGTCA 1950
Db |||||

409 CAGACTTTACCGCCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGGATCACACAGAAAC 468
Db |
622 CAGACTTTACCGCCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGGATCACACAGAAAC 681
Qy |
469 TCCTCA-----GTGGCCCTTTCAACCTGCCAGGCTTGTC 504
Db ||
682 TGCCATCGCATGGAGGGAAGTTCTGTGAGGGCTCCACTCGCACTCTGAAGCTCTGCAAC 741
Qy ||
505 GGATGATACGAACAGAGGAGGAGATTAATCTTCTAAGGCCACTTCTTTCACACCTCTCA 564
Db ||
742 AGTCAGAAATGTCCCGGACAGTGTGACTTCCGTGCTCAGTGTGCGAGCACAAC 801
Qy ||
565 TGGAACTCGGAGAGCTGCCAAGGACGCTGCCATCCACGTACTGTACAGAGATCC 624
Db ||
802 AGCAGACGATTCAGAGGGCGGCATCAAGTGG--AAGCCTTACACTCAAGTAGAGCCG 859
Qy ||
625 ACAGAGCCCATGCTCCTGGGGCCAGTGAGGTCCTGGTGACCTC-----AAGG 672
Db ||
860 ACTTATGCAAACTCTACTGTATCGCAGAAGGATTTGATTTCTTCTTTTCTTTGTCAATA 919
Qy ||
673 ACATGGAGCTGGCAATCAACCCCTGCAAGAGCGACCTTGGCCCTGGGACTGCCACAA 732
Db ||
920 AAGTCAAGATGGGACTCCATGCTCGAGGATAGCCGTAATGTTTGTATAGATGGGATAT 979
Qy ||
733 AAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCGCCAGCTCCCAAGGAGACCTC 792
Db ||
980 GTGAGCTCAGTGTGGTGTCCACATCTG-CGCACATGCCCGCAGCTCCCAAGGAGACCTC 1038
Qy ||
793 TTCTATCTGCCAGATGAGTATAGTCTTCTTACGGCATAAGCGCTCTCTCTGAGGTCC 852
Db ||
1039 TTCTATCTGCCAGATGAGTATAGTCTTCTTACGGCATAAGCGCTCTCTCTGAGGTCC 1098
Qy ||
853 CATAGAAATGAAGAACTGAACGTGGAGACTTGTGGTGGTGGTGGACAAAGATGATGCAA 912
Db ||
1099 CATAGAAATGAAGAACTGAACGTGGAGACTTGTGGTGGTGGTGGACAAAGATGATGCAA 1158
Qy ||
913 AACCATGGCCATGAATAATACACCACTAGCTGCTCAGATACTCAATGATGATCTGCT 972
Db ||
1159 AACCATGGCCATGAATAATACACCACTAGCTGCTCAGATACTCAATGATGATCTGCT 1218
Qy ||
973 TTATTCAAAGATGGAACATATAGGAGGAACATCAATTGCAATTGTAGGTCTGATCTT 1032
Db ||
1219 TTATTCAAAGA-----1229
Qy ||
1033 CTAGAAGATGAACAGCCAGGACTGGTGATAAGTCAACACGACAGACCACTTAAAGTAGC 1092
Db ||
1230 -----1229
Qy ||
1093 TTCTGCGAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCAACACCGCCATC 1152
Db ||
1230 -----TGGATTGATGGGAAAGATGGGACTCGTCAACACCGCCATC 1272
Qy ||
1153 TTACTGACTGCTGGATATATGTTCTCGAAGATGAGCCCTGTGACACTTTGGGATTT 1212
Db ||
1273 TTACTGACTGCTGGATATATGTTCTCGAAGATGAGCCCTGTGACACTTTGGGATTT 1332
Qy ||
1213 GCACCCATAAGTGGATGTAGTAAATATCGCAGCTGCACGATTAATGAAGATACAGGT 1272
Db ||
1333 GCACCCATAAGTGGATGTAGTAAATATCGCAGCTGCACGATTAATGAAGATACAGGT 1392
Qy ||
1273 CTGGAATGGCTTCAACCATGCCATAGTCTGGACAAATCTTGGGATGATCATGAT 1332
Db ||
1393 CTGGAATGGCTTCAACCATGCCATAGTCTGGACAAATCTTGGGATGATCATGAT 1452
Qy ||
1333 GGAGAGGGAACATGTGTAAGTCCGAGGCAACATCATGTCCTCCATCATTTGGCAGGA 1392
Db ||
1453 GGAGAGGGAACATGTGTAAGTCCGAGGCAACATCATGTCCTCCATCATTTGGCAGGA 1512
Qy ||
1393 CGCAATGGAGTCTTCTCTGCTCACCCTGAGCCGAGTATCTACACAAATTTCTAAGC 1452
Db ||
1513 CGCAATGGAGTCTTCTCTGCTACCCCTGAGCCGAGTATCTACACAAATTTCTAAGC 1572
Qy ||
1453 ACCGCTCAAGCTATCTGCTTGTGATCAGCCAAAGCCCTGTGAAGGAATACAAAGTATCCT 1512

1573 ACCGCTCAAGCTATCTGCTTGTGATCAGCCAAAGCCCTGTGAAGGAATACAAAGTATCCT 1632
Qy ||
1513 GAGAAATTCGCCAGGAGAAATATATGATGCAAAACACACAGTGCAGTGGCAGTTTCGAGAG 1572
Db ||
1633 GAGAAATTCGCCAGGAGAAATATATGATGCAAAACACACAGTGCAGTGGCAGTTTCGAGAG 1692
Qy ||
1573 AAAGCCAAGCTCTGCATGTGCACTTTAAAAAGGACATCTGTAAAGCCCTCTGTGTCCTAT 1632
Db ||
1693 AAAGCCAAGCTCTGCATGTGCACTTTAAAAAGGACATCTGTAAAGCCCTCTGTGTCCTAT 1752
Qy ||
1633 CGTATTGGAGGAAATGTGAGACTAAATTTATGCCAGCAGCAGAGGCAATTTGTGGG 1692
Db ||
1753 CGTATTGGAGGAAATGTGAGACTAAATTTATGCCAGCAGCAGAGGCAATTTGTGGG 1812
Qy ||
1693 CATGACATGTGTGCGGGGAGGACAGTGTGTGAATAATGTGTGATGAAGGCCCAAGCCCC 1752
Db ||
1813 CATGACATGTGTGCGGGGAGGACAGTGTGTGAATAATGTGTGATGAAGGCCCAAGCCCC 1872
Qy ||
1753 ACCCATGGCCACTGTGTGGAAGTGGTCTTCTGGTCCCCATGTCTCAGGACCTGCGAGGG 1812
Db ||
1873 ACCCATGGCCACTGTGTGGAAGTGGTCTTCTGGTCCCCATGTCTCAGGACCTGCGAGGG 1932
Qy ||
1813 GGAGTATCTCATAGGAGTGCCTCTGCACCAACCCCAAGCCCATGCGATGGAGGAAAGTTC 1872
Db ||
1933 GGAGTATCTCATAGGAGTGCCTCTGCACCAACCCCAAGCCCATGCGATGGAGGAAAGTTC 1992
Qy ||
1873 TGTGAGGGTCCCACTCGCACTCTGAAGCTCTGCACACAGTCAAGAAATGTCCCGGACAGT 1932
Db ||
1993 TGTGAGGGTCCCACTCGCACTCTGAAGCTCTGCACACAGTCAAGAAATGTCCCGGACAGT 2052
Qy ||
1933 GTTGACTTCCGCTGCTCAGTGTGCGAGCAACACAGCAGACGATTCAGAGGCGCGCAC 1992
Db ||
2053 GTTGACTTCCGCTGCTCAGTGTGCGAGCAACACAGCAGACGATTCAGAGGCGCGCAC 2112
Qy ||
1993 TACAAGTGAAGCCTTACACTCAAGTAGAAGTCAAGGACTTATGCAAACTCTACTGTATC 2052
Db ||
2113 TACAAGTGAAGCCTTACACTCAAGTAGAAGTCAAGGACTTATGCAAACTCTACTGTATC 2172
Qy ||
2053 CGAGAAGGATTTGATTTCTTTCTTTGTCAAATAAAGTCAAGATGGGACTCCATGTC 2112
Db ||
2173 GCAGAAGGATTTGATTTCTTTCTTTGTCAAATAAAGTCAAGATGGGACTCCATGTC 2232
Qy ||
2113 TCGGAGGATAGCCGTAATGTTTGTATAGATGGGATAATGTGAGAGAGTTGGATGTGCAAT 2172
Db ||
2233 TCGGAGGATAGCCGTAATGTTTGTATAGATGGGATAATGTGAGAGAGTTGGATGTGCAAT 2292
Qy ||
2173 GTCTTTGGATCTGATGCTGTTGAAGACGCTCTGTGGGGTGTAAACGGGAATTAACCTCAGCC 2232
Db ||
2293 GTCTTTGGATCTGATGCTGTTGAAGACGCTCTGTGGGGTGTAAACGGGAATTAACCTCAGCC 2352
Qy ||
2233 TGCACGATTTACAGGGGTCTCTACACCAAGCACCACCAACAGTATTAATCAATG 2292
Db ||
2353 TGCACGATTTACAGGGGTCTCTACACCAAGCACCACCAACAGTATTAATCAATG 2412
Qy ||
2293 GTCCACATTTCTTTGGAGCCCGGAGTATCCGATCTATGAATAAGACCTCTACTCCTCC 2352
Db ||
2413 GTCCACATTTCTTTGGAGCCCGGAGTATCCGATCTATGAATAAGACCTCTACTCCTCC 2472
Qy ||
2353 TACATTTCTGTGCGCAATGCCCTCAGAAAGTACTACTCTGAAATGGGCACCTGGACCGTGGAC 2412
Db ||
2473 TACATTTCTGTGCGCAATGCCCTCAGAAAGTACTACTCTGAAATGGGCACCTGGACCGTGGAC 2532
Qy ||
2413 TGGCCCGCCCGGTACAAATTTTCGGGCACTACTTTCGACTACAGACGGTCTTATATAG 2472
Db ||
2533 TGGCCCGCCCGGTACAAATTTTCGGGCACTACTTTCGACTACAGACGGTCTTATATAG 2592
Qy ||
2473 CCCGAGAACTTAATCGCTACTGGAACCAACAGAGACACTCATTTGTGGAGCTGCTGTTT 2532
Db ||
2593 CCCGAGAACTTAATCGCTACTGGAACCAACAGAGACACTCATTTGTGGAGCTGCTGTTT 2652
Qy ||
2533 CAGGAAAGGAACCCGGGTGTCCTGGGAATACTCAATGCCCTGCTTGGGGAACCGGAAAG 2592

```
Db      2653 CAGGGAAGAACCCGGGTGTTGCTTGGGAATACCTCCATGCTCCTTTGGGGAACGAGAAG 2712
QY      2593 CAGCCCCCTGCCAGCCAGCTACACTTTGGGCCATCGTGGCTCTGAGTGTCTCGGTGCC 2652
Db      2713 CAGCCCCCTGCCAGCCAGCTACACTTTGGGCCATCGTGGCTCTGAGTGTCTCGGTGCC 2772
QY      2653 TGGCGAGGGGG 2663
Db      2773 TGGCGAGGGGG 2783

RESULT 12
US-10-217-774-1
; Sequence 1, Application US/1021774
; Publication No. US2002019358A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. US20020193583A1el Human Proteases and Polynucleotides Encodir
; FILE REFERENCE: LEX-0219-USA
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US/09/930,872
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1476
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-217-774-1

Query Match      39.5%; Score 1449.8; DB 13; Length 1476;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1454; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      1  ATGAAGCCCGCGCGCGATGCGGGGCTTGGCGGCTCTGGATGCTGCTGGCGGAG 60
Db      1  ATGAAGCCCGCGCGCGATGCGGGGCTTGGCGGCTCTGGATGCTGTTGGCGGAG 60

QY      61  GTGGCGAGCAGGCACTCGTGCCTATGGACCGCGAGCGCGAGCGCTGGGAGCCG 120
Db      61  GTGGCGAGCAGGCACTCGTGCCTATGGACCGCGAGCGCGAGCGCTGGGAGCCG 120

QY      121 AGCGTCCCGCTCTCTCCACCGCGAGCGCGGCTGGATGGAAAGGCGAATAT 180
Db      121 AGCGTCCCGCTCTCTCTCCACCGCGAGCGCGGCTGGATGGAAAGGCGAATAT 180

QY      181 GACCTGGTCTCTGCTTACGAGTTGACACAGGGGCGATTACGTGTCCTCATGAATCATG 240
Db      181 GACCTGGTCTCTGCTTACGAGTTGACACAGGGGCGATTACGTGTCCTCATGAATCATG 240

QY      241 CACCATCAGCGCGGAGAGAGAGTGGCCGTGTCGAGGTTGAGTCTTTCACTTCGG 300
Db      241 CACCATCAGCGCGGAGAGAGAGTGGCCGTGTCGAGGTTGAGTCTTTCACTTCGG 300

QY      301 CTGAAAGGCTCCAGGCAAGAGTGGGAAAGACAGGCACTAAGTCTGTGCAGACTTTACCG 420
Db      301 CTGAAAGGCTCCAGGCAAGAGTGGGAAAGACAGGCACTAAGTCTGTGCAGACTTTACCG 420

QY      361 CTTGGTTTATTTGTGAGAGCTTTGAGGAAAGACAGGCACTAAGTCTGTGCAGACTTTACCG 480
Db      361 CTTGGTTTATTTGTGAGAGCTTTGAGGAAAGACAGGCACTAAGTCTGTGCAGACTTTACCG 480

QY      481 CTTTCAACTGCCAAGGCTTGTGAGGCAATGATCGAAACAGAAAGGCGAGATTACTTCTTA 540
Db      481 CTTTCAACTGCCAAGGCTTGTGAGGCAATGATCGAAACAGAAAGGCGAGATTACTTCTTA 540
```

```
Db      481 CTTTCAACTGCCAAGGCTTGTGAGGCAATGATCGAAACAGAAAGGCGAGATTACTTCTTA 540
QY      541 AGGCCACTTCTTCAACACTCTCATTGGAAACTCGGAGAGCTGCCAAGGCGAGCTCGCCA 600
Db      541 AGGCCACTTCTTCAACACTCTCATTGGAAACTCGGAGAGCTGCCAAGGCGAGCTCGCCA 600
QY      601 TCCCACTGACTGTACAGAGATCCACAGAGCCCATGTCTCTGGGGCCAGTGAGGTCCTG 660
Db      601 TCCCACTGACTGTACAGAGATCCACAGAGCCCATGTCTCTGGGGCCAGTGAGGTCCTG 660
QY      661 GTGACCTCAAGGACATGGGAGCTGGCAATCAACCCCTGCAAGAGAGGAGCTTCCGCTG 720
Db      661 GTGACCTCAAGGACATGGGAGCTGGCAATCAACCCCTGCAAGAGAGGAGCTTCCGCTG 720
QY      721 GGACTGCCACAAAAGCAGCATTTCTGTGGAAGCGCAAGAAATACATGCCCGAGCTGCC 780
Db      721 GGACTGCCACAAAAGCAGCATTTCTGTGGAAGCGCAAGAAATACATGCCCGAGCTGCC 780
QY      781 AAGGAAGACCTCTTTCATCTTCCAGATGAGTATAGTCTTGTACGGCATAAAGGCTCT 840
Db      781 AAGGAAGACCTCTTTCATCTTCCAGATGAGTATAGTCTTGTACGGCATAAAGGCTCT 840
QY      841 CTTCTGAGGTCCCATAGAAATGAAGACTGAACGTGGAGACTTGGTGGTGGTGGTGGTGG 900
Db      841 CTTCTGAGGTCCCATAGAAATGAAGACTGAACGTGGAGACTTGGTGGTGGTGGTGGTGG 900
QY      901 AAGATGATGCAAAAACCATGGCCATCAAAATATCACCACTACGTGCTCAGCATCTCAAC 960
Db      901 AAGATGATGCAAAAACCATGGCCATCAAAATATCACCACTACGTGCTCAGCATCTCAAC 960
QY      961 ATGTTATCTGCTTTATTTCAAGATGGAACATAGAGGAAACATCAACATTTGCAATTGTA 1020
Db      961 ATGTTATCTGCTTTATTTCAAGATGGAACATAGAGGAAACATCAACATTTGCAATTGTA 1020
QY      1021 GGTCTGATTTCTTAGAAGATGAACAGCAGGACTGGTGTATAGTCAACCGAGAGCAC 1080
Db      1021 GGTCTGATTTCTTAGAAGATGAACAGCAGGACTGGTGTATAGTCAACCGAGAGCAC 1080
QY      1081 ACCTTAAGTACTTCTCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCAAT 1140
Db      1081 ACCTTAAGTACTTCTCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCAAT 1140
QY      1141 GACCACGCCATCTTACTGACTGGTCTGGATATATGTTCTTGGAAAGATGAGCCCTGTGAC 1200
Db      1141 GACCACGCCATCTTACTGACTGGTCTGGATATATGTTCTTGGAAAGATGAGCCCTGTGAC 1200
QY      1201 ACTTTGGGATTTGACCCATAAGTGGAAATGTAGTAAATATCGCAGCTGCAAGATTAAAT 1260
Db      1201 ACTTTGGGATTTGACCCATAAGTGGAAATGTAGTAAATATCGCAGCTGCAAGATTAAAT 1260
QY      1261 GAAGATACAGGTCTTGGACTGGGCTTCAACATTTGCCCATGAGTCTGGACACAATTTGGC 1320
Db      1261 GAAGATACAGGTCTTGGACTGGGCTTCAACATTTGCCCATGAGTCTGGACACAATTTGGC 1320
QY      1321 ATGATTCATGATGGAGAGGAAACATGTGTAAAGTCCGAGGGCAACATCATGTCCCT 1380
Db      1321 ATGATTCATGATGGAGAGGAAACATGTGTAAAGTCCGAGGGCAACATCATGTCCCT 1380
QY      1381 ACATTTGGCAGGACGCAATGGAGTCTTCTCTGTGTCACTCCCTGCGAGCGCGAGTATCTACAC 1440
Db      1381 ACATTTGGCAGGACGCAATGGAGTCTTCTCTGTGTCACTCCCTGCGAGCGCGAGTATCTACAC 1440
QY      1441 AAATTTCTAAGCACCCCTCAA 1461
Db      1441 AAATTTCTAAGTCACTGAA 1461
```

```
RESULT 13
US-10-804-457-1
; Sequence 1, Application US/10804457
; Publication No. US2005006534A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
```

APPLICANT: Hilbun, Erin
TITLE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the
FILE REFERENCE: LEX-0219-USA
CURRENT APPLICATION NUMBER: US/10/804,457
CURRENT FILING DATE: 2004-03-19
PRIOR APPLICATION NUMBER: US/10/217,774
PRIOR FILING DATE: 2002-08-12
PRIOR APPLICATION NUMBER: US/09/930,872
PRIOR FILING DATE: 2001-08-14
PRIOR APPLICATION NUMBER: US 60/225,852
PRIOR FILING DATE: 2000-08-16
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1476
TYPE: DNA
ORGANISM: homo sapiens
US-10-804-457-1

Query Match 39.5%; Score 1449.8; DB 21; Length 1476;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1454; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY	1	ATGAAGCCCGCGCGCGGATGGCGGCTTGGCGCGCTGTGGATGCTGTGGCGCAG	60
DB	1	ATGAAGCCCGCGCGCGGATGGCGGCTTGGCGCGCTGTGGATGCTGTGGCGCAG	60
QY	61	GTGGCGGAGCAGCAGCCTGGCTGGCGCATGGGACCGCGCGCGAGCGGCTGGGAGCCCG	120
DB	61	GTGGCGGAGCAGCAGCCTGGCTGGCGCATGGGACCGCGCGCGAGCGGCTGGGAGCCCG	120
QY	121	AGCGTCCCGCGTCTCTCCACCGCGGAGCGCGCGGCTGGATGGAAAAGGCGCAATAT	180
DB	121	AGCGTCCCGCGTCTCTCCACCGCGGAGCGCGCGGCTGGATGGAAAAGGCGCAATAT	180
QY	181	GACCTGGTCTCTGCTACGAGGTGTGACCAAGGGGCGATTAAGTGTCCCATGAATCATG	240
DB	181	GACCTGGTCTCTGCTACGAGGTGTGACCAAGGGGCGATTAAGTGTCCCATGAATCATG	240
QY	241	CACCATCAGCGCGGAGAGAGCAGTGGCGGTCCGAGGTTGAGTCTCTTCACTTCGG	300
DB	241	CACCATCAGCGCGGAGAGAGCAGTGGCGGTCCGAGGTTGAGTCTCTTCACTTCGG	300
QY	301	CTGAAGAGGCTCCAGGCAAGCTTCCAGTGGATCTCAGGAGCTTCCAGAGCCCTAGTGGCT	360
DB	301	CTGAAGAGGCTCCAGGCAAGCTTCCAGTGGATCTCAGGAGCTTCCAGAGCCCTAGTGGCT	360
QY	361	CTGGGCTTTATGTGCAGAGTTGGGAAAGACAGGCACTAAGTCTGTGCAGACTTTACCG	420
DB	361	CTGGGCTTTATGTGCAGAGTTGGGAAAGACAGGCACTAAGTCTGTGCAGACTTTACCG	420
QY	421	CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGGATCACAAGAACTCTCAGTGGCC	480
DB	421	CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGGATCACAAGAACTCTCAGTGGCC	480
QY	481	CTTTTCAACCTGCCAAGGCTTGTCCAGCATGATACGAAACAGAGAGGAGGATCTTCTTA	540
DB	481	CTTTTCAACCTGCCAAGGCTTGTCCAGCATGATACGAAACAGAGAGGAGGAGGATCTTCTTA	540
QY	541	AGGCCACTTCTTTTCAACCTCTCATGAAACTCCGAGAGCTGCCCCAAGGCACTCGCCA	600
DB	541	AGGCCACTTCTTTTCAACCTCTCATGAAACTCCGAGAGCTGCCCCAAGGCACTCGCCA	600
QY	601	TCCACGTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	660
DB	601	TCCACGTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	660
QY	661	GTGACCTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	720
DB	661	GTGACCTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	720
QY	721	GGACTGCCCAAAAGCAGCATTTTCTGTGGAGAGCGCAAGAAATPACATGCCCGCCTCCC	780

DB	721	GGACTGCCCAAAAGCAGCATTTTCTGTGAAGACGCAAGAAATACATGCCCGCCTCCC	780
QY	781	AAGGAAGACCTCTTCTATCTTCCAGATGATTAAGTCTTCTTACGGATAAGCGCTCT	840
DB	781	AAGGAAGACCTCTTCTATCTTCCAGATGATTAAGTCTTCTTACGGATAAGCGCTCT	840
QY	841	CTTCTGAGGTCCCATAGAAATGAAGAACTGAACGTGGAGACCTTGGTGGTGGTGGTGG	900
DB	841	CTTCTGAGGTCCCATAGAAATGAAGAACTGAACGTGGAGACCTTGGTGGTGGTGGTGG	900
QY	901	AGATGATGCAAAACCATGCGCATGAAATATCACCACTACCTGCTCAGGATCTCAAC	960
DB	901	AGATGATGCAAAACCATGCGCATGAAATATCACCACTACCTGCTCAGGATCTCAAC	960
QY	961	ATGGTATCTGCTTTTATTTCAAGATGCAACAATAGGAGGAAACATCAACATTTGCAAT	1020
DB	961	ATGGTATCTGCTTTTATTTCAAGATGCAACAATAGGAGGAAACATCAACATTTGCAAT	1020
QY	1021	GGTCTGATTTCTTAGAAGATGAACAGCAGGACTGCTGATTAAGTCAACCGCAGACCC	1080
DB	1021	GGTCTGATTTCTTAGAAGATGAACAGCAGGACTGCTGATTAAGTCAACCGCAGACCC	1080
QY	1081	ACCTTAAGTACCTTCTGCGAGTGGGAGTCTGATTAAGTGGGAGAAAGATGGGACTCGT	1140
DB	1081	ACCTTAAGTACCTTCTGCGAGTGGGAGTCTGATTAAGTGGGAGAAAGATGGGACTCGT	1140
QY	1141	GACCAAGCCTCTTCTGAGTCTGCTGATTAAGTGGGAGAAAGATGGGACTCGT	1200
DB	1141	GACCAAGCCTCTTCTGAGTCTGCTGATTAAGTGGGAGAAAGATGGGACTCGT	1200
QY	1201	ACTTTGGGATTTGCACCCATTAAGTGGAAATGTGTAGTAAATATCGCAGCTGCAAGTAA	1260
DB	1201	ACTTTGGGATTTGCACCCATTAAGTGGAAATGTGTAGTAAATATCGCAGCTGCAAGTAA	1260
QY	1261	GAAGATACAGGCTTCTGAGTGGGCTTCCACATTTGCCCATGAGTCTGGAACAACCTTGG	1320
DB	1261	GAAGATACAGGCTTCTGAGTGGGCTTCCACATTTGCCCATGAGTCTGGAACAACCTTGG	1320
QY	1321	ATGATTTATGATGAGGAGGAGGAAACATGTCTAAAGTCCGAGGCAACATCATGTCCCT	1380
DB	1321	ATGATTTATGATGAGGAGGAGGAAACATGTCTAAAGTCCGAGGCAACATCATGTCCCT	1380
QY	1381	ACATTTGGCAGGAGCGCAATGAGTCTTCTCTGCTACCCCTGCGAGCGCCAGTATCTAC	1440
DB	1381	ACATTTGGCAGGAGCGCAATGAGTCTTCTCTGCTACCCCTGCGAGCGCCAGTATCTAC	1440
QY	1441	AAATTTCTAAGCACCGCTCAA	1461
DB	1441	AAATTTCTAAGCATCAGTAAA	1461

RESULT 14
US-10-240-545A-1
Sequence 1, Application US/10240545A
Publication No. US20030185828A1
GENERAL INFORMATION:
APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
TITLE OF INVENTION: No. US20030185828A1el aggreganase
FILE REFERENCE: 08959.0002
CURRENT APPLICATION NUMBER: US/10/240,545A
CURRENT FILING DATE: 2002-10-02
PRIOR APPLICATION NUMBER: PCT/JP01/11033
PRIOR FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: JP 2000-384300
PRIOR FILING DATE: 2000-12-18
NUMBER OF SEQ ID NOS: 26
SEQ ID NO 1
LENGTH: 3666
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS

; LOCATION: (1)..(3666)
US-10-240-545A-1

	Query Match	31.0%;	Score 1139.4;	DB 16;	Length 3666;
	Best Local Similarity	60.1%;	Pred. No. 0;	Mismatches 1341;	Indels 69; Gaps 11;
	Matches 2123;	Conservative 0;			
QY	168	AAMGGGCGAATATCAGCTGTGTCCTGCTACGAGGTTGACCACAGGGGGCATTTACGTGC	227		
DB	171	AAATGATGATTACGTCTTTGTACGCCAGTAGAAGTAGACTTCAGCCGGGTCAATATAATTC	230		
QY	228	CCATGAANAATCATGCACCATCAGCGCGGAGAAGACAGTGGCCGTGCCAGGTTTGAGTC	287		
DB	231	ACACGACATTTTGCAC---AACGCGACGAAAAAGCGATCGSCGCAGAAATCCAGAACTC	287		
QY	288	TCTTTCACCTTCGGCTGAAGAGCTCCAGGCACGACTCCACGTGGATCTGAGGACTCCAG	347		
DB	288	CCTGCACACTCGGATTTACATTTTGACACGGAATGCACTTAAGACTTAAGCCCTCGGC	347		
QY	348	CAGCCTAGTGGCTCCTGGCTTTATFTGTGCAGACGTTGGGAAAGACAGGCACTAAGTCTGT	407		
DB	348	GA---TTTTGAGCAGTCACTTTATTTGTCCAGGTACTTGGAAAAGATGGTCT---TCAGA	401		
QY	408	GCAGACTTTTACGCCACAGAGACTTCTGTTTCTATCAAGGCTCTTTGGATFCAACAAGAA	467		
DB	402	GACTCAGAAAACCGAGGTGCAGCAATGCTTCTATCAGGGATTTTATCAGAATCACAGCTC	461		
QY	468	CTCCTCAGTGGCCCTTTCAAACCTGCCAAGGCTTGTGAGGCATCATACGAACAGAGGC	527		
DB	462	CTCCTCTGTGCTGTGTACGTGTGCTGSGTGTTCAGGTTTAATTAAGNACAGAAAATA	521		
QY	528	AGATTACTTCCTAAGGCGCACTTCTTTCACACCTCTCATGGAAACTCGCGAGAGCTGCCA	587		
DB	522	TGAATTCCTCATCTGCCAATTAACCTCAGCTTCTGGCCCCAGGAACAACAATACAGCTCCC	581		
QY	588	AGCAGCTGCCATCCACGTACTGTACAAGAGATCCACAGAGCCCCATGCTCTCGGGGC	647		
DB	582	TGCGGTTCACCATCTCTCAGTACTGTACAAAAGGACAGCAGAGGAGAAGTCCAGCGTA	641		
QY	648	CAGTAGGTCCTGGTGACCTCAAGGA-----CATGGGAGCTGGCAATCAACCCCTGCAAC	702		
DB	642	CCGTGGCTACCCCGGCTCTGGCCGGAAATTTATCTGGTTACTCCCCAAGTCACATTCCCA	701		
QY	703	AGCAGC-----GACCTTCGCTGGGACTGCCCCACAAAAAGCAGCATTTCTG	746		
DB	702	TGCATCTCAGAGTCGAGACACAGAGTATCAACCATCGAAGTTGCAAAAGCAGCATTTTGT	761		
QY	747	TGSRAGACGCAAGAAATACATGCCCCAGCTCCCAAGGAAGACTCTTCATCTTCCAGAA	806		
DB	762	TGGACGACGCAAGAAATATGCTCCCAAGCTCCACAGAGGACACCTATCTAAGGTTTGA	821		
QY	807	TGAGTATAAGTCTTGCTTACGGCATAAAGCGCTCTCTTCTGAGGTCCCATAGAAATGAAGA	866		
DB	822	TGAATATGGAGCTCTGGGCGACCCAGAGATCAGCTGGAAATCACAA-----AGGG	875		
QY	867	ACTGAACTGGAGACTTGGTGGTGTGACAAAAAGATGATGCAAAACCATGGCCATGA	926		
DB	876	CCTCAATGTGGAAAACCTCGTGGTGGCAGACAAGAAATGGTGGAAAAAGCATGGCAAGGG	935		
QY	927	AAATATACCACCTACGTGCTCACCATACTCAACATGGTATCTGCTTTATTTCAGAAGTGG	986		
DB	936	AAATGTCACCACATACATTTCTCACAGTAATGACATGGTTTCTGGCTATTTTAAAGATGG	995		
QY	987	AACRATAGGAGGAACATCAATTTGCAATTTGTAGGTCTGATTTCTTCTAGAGATGAACA	1046		
DB	996	GACTATTGGAAAGTGACATAAAGCGTGGTTGTGGTGAAGCCCTAATTTCTTCGAAACAGAACC	1055		
QY	1047	GCCAGGACTGGTGAATGATCAACGCGCAGACACACCTTTAAGTAGCTTCTGCCAGTGCA	1106		
DB	1056	TGGAGGATTAATTGATCAACCATCATGCGACCAAGTCTCTGAATAGTTTGTCTCAATGGCA	1115		
QY	1107	GTCTGGATTGATGGGAAAGATGGGACTCGTFCATGACCAAGCCATCTTACTGACTGGTCT	1166		


```
QY 2247 GGCTCTTACCAAGGACCAACACAGCAACAGTATTATATCATATGTCACCATTCCTTC 2306
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2256 AGGCTGTATCCTCAACAGCAGATGAATGAATGATATATTCGGTGGTCAATTCACG 2315
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2307 TGGAGCCCGAGTATCGCATCTATGAATGAACGCTCTTACCTCTTACATTTCTGTGG 2366
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2316 TGGCGCCGAGATGAAATCCAGAGCTGACGTTTCTCAGTTACCTCGCATTCG 2375
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2367 CAATGCCCTCAGAGGTACTTACCTGAATGGGCACTGGACCGTGGACTGGCCCGCGGTA 2426
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2376 AAGCCTCAGTCAAAGATATTACCTCACCGGGGCTGGAGCATCGACTGGCTGGGAGTT 2435
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2427 CAATTTTTCGGGACATCTTTCAGCTACAGAGGCTCTATATGAGCCCGAGAACCTTAAT 2486
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2436 CCCCTTCGCTGGGACACGTTTGAATACCAAGCGCTCTTCAACCGCCCGGAACGCTGTA 2495
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2487 CGCTACTGGACCAACCAACGAGACACTGATTGGAGCTGCTGTTTCAGGAGGAACCC 2546
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2496 CGGCCAGGCCCCACAAATGAGACGCTGGTCTTTGAAATTCGATCAAGGCAAAATCC 2555
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2547 GGCTGTTCCTGGGAATACTCCATGCTCTCGCTTGGGGACCGAGAAGCAGCCCCCTGCCCA 2606
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2556 AGGATAGCTTGAAGTATCCACTTCCCAAGTCAATGAATGGAATCCACGCCACAAA 2615
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2607 ---GCCAGCTACATTGGGCCATCGTGGCTCTGAGTGTCTCGTCTCTGGGAGGGG 2663
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2616 AAGACCTGCTATACCTGGAGTATCGTGAGTCAGAGTGTCTCGTCTCTGTGTGGAGG 2675
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2664 ACAGATGACGCTGAGAGGGCTGCTACAGAGACCTGAAGTTCAAGTAATATGCTT 2723
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2676 TTACATAAATGTAAAGGCCATTGGCTTGGAGATCAAAATACTCAAGTCAATTCCTCAT 2735
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2724 CTGCAATCCCAAGACACGACTCTGACGGGGCTGGTCTGCAAAAGTATCTGCCGTGC 2783
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2736 CTGCAATGCAAAACCAAGCAGTAACTGAGCCCAAAATCTGACGCTTCTCTGCC 2795
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2784 TCCAGCTGGTCCGTGGGAACTGGAGTGCCTGAGTCCGAGTCCGAGCGGGGGTCCCA 2843
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2796 GGCTTACTGGTCCAGGTGMAATGGAGTACATGCAAGGCGCTGTGCTGGAGCCAGCA 2855
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2844 GAGCGCCCGTGCAGTGCACAGCGGGGTGCACT---ATGACTCGAGCCAGTCCCGC 2900
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2856 GAGCGAAAGATCCAGTGTGTGCAAAAGAGCCCTTCCAAAAGGAGGACGATGTGCA 2915
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2901 CAGCTGTGCTCTCAGCTGCTCCCTCCAGCAGCAGGCTGCAACTCTCAGAGCTGCC 2960
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2916 TTCTCTCTCAGTGCAGCACCCCACTCAGGTCCAGCTGCAAGCAGCATCTCTGCC 2975
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2961 ACTGTCATGAGCGCCCGGCTGGGACAGTGTCTCACAACATGTGGGAAGGGGTGGAG 3020
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2976 TCCACAATGAGCGCTTGGACCCCTGGTCTCAGTGTCCAAAGACCTGTGGACGAGGTGAG 3035
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3021 GAGCGGGCAGTGGCTGTAAAGAGCACAAACCCCTCGGCCAGAGCGCAGCTGTGCCCA 3080
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3036 GAAGCGTGAATCTCTCTGAAGGGCTCTGCC-----GCAGAAACCCCTCCCGA 3083
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3081 CGCTGTCTCAGCTCCGAGCCCAAGCCAGAGTGCATGAGCTGTCTCTCAGCGCTG 3140
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3084 GAGCAGTGTACAGTCTCCCAAGCTGAGCTGACGAGGGCTGTGTCTTGGAGGATG 3143
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3141 CCAAGGCCCAAGAGCTGAGTGGTGGTGTCCGCTCTGGTCCAGTGTCTGTGACATG 3200
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3144 CCCCAAGACAGCGGCTACAGTGGTCTCTTCTGTGGAGGAGTGTCTGCAACCTG 3203
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3201 TGAAGAGGAAACAGAAAGATTTCTAAATGTGCTGAAAGTATGTTTCTGGAAGTA 3260
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3204 TGGTTTGGTGTGAGAGAGGGAGATGAAGTGCAGCAGAGAGGGCTTCCAGGGAAGCT 3263
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3261 TCGAGAGCTGGCTCAAGAGTGTCTCATATTTGCCGAAGCCAGCGCTGAGCTGAACG 3320
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3264 GATACTTTCCAGAGCGAAGATGCCGTAAATATTAAAGAAACCAATCTGGACTTGAAGA 3323
```

```
QY 3321 TGCCTGGCCCGCTTCCATTCATGCCCCAGGCAACCCCAATTTGTGCTGGGACCTCGAG 3380
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3324 GACCTGCAACCCAGCGGCTTGCCAGCCCA-----TCCAGTGTACAAACATGGT 3371
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3381 GGGCAGCTGTTTGGCTCACCTCTCAGTGCAGCGCAGCTGTGGGAGGCGTTCA 3440
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3372 AGCTGATGATTTATTCCTCCGCTGGCAGCAGTGCACGTCTGTGGGAGGGGTCCA 3431
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3441 GAGCAGGTCCTGTGCACTGTGCTGGTGGGCGCGGCTCAGGCTGCTCTTGCACCA 3500
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3432 GACCCGGTCAGTCCACTGTGTTTTCAGCAAGCGCGCTTCTCAAGTTGTCTGCTCATCA 3491
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3501 GAGCCTTCGGGCTCCCTGGCTGCAACACTCACTCTTGTGCCCCATTCGACAGAGA---A 3557
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3492 GAAACCTCCGGTGTACGAGCGCTGTAATAACAACTTCTGTCCAGCTCTGTAAGAGAGA 3551
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3558 AGATGCTTCTGCAAAAGACTTACTTCCACTGTGCTACCTGGTACCCAGCGGGATGTG 3617
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3552 GGATCATCTCGGTAGATTTCTTCACTGGTGTACCTAGTTCCTCAGCATGGTGTCTG 3611
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3618 CAGCCCAAGTTTCTAGGCAAGCAGTGTGCAAGACTTGTCTTAAGTCCAACT 3670
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3612 CAACACAGTTTTCGGGAAACAATGCTGCAAGTCAATGCACAAGGAAGATCT 3664

RESULT 15
US-10-354-983-1
; Sequence 1, Application US/10354983
; Publication No. US20040044194A1
; GENERAL INFORMATION:
; APPLICANT: AGOSTINO, MICHAEL J.
; APPLICANT: CORCORAN, CHRISTOPHER
; TITLE OF INVENTION: AGGRECANASE MOLECULES
; FILE REFERENCE: 08702.0111-00000
; CURRENT APPLICATION NUMBER: US/10/354,983
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: 60/353,680
; PRIOR FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3663
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-354-983-1

Query Match 30.9%; Score 1135.6; DB 18; Length 3663;
Best Local Similarity 60.1%; Pred. No. 0;
Matches 2118; Conservative 0; Mismatches 1339; Indels 69; Gaps 11;

QY 168 AAAGGGCGGAATATGACCTGGTCTCTGCTTACGAGGTGTGACACAGGGGCGGATTTACGTGTC 227
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
171 AAATGATGATTACGCTTTTGTTCACGCCAGTAGAGTAGACTCAGCGGGGTCAATATTTTC 230
QY 228 CCATGAATATCATGCCACATCAGCGCGGAGAGAGACAGTGGCGGTGTCCGAGGTTGAGTC 287
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
231 ACACGACATTTTGCAC---AAGCGCAGGAAAGCGATCGGCGCAGAAATGCCAAGAGCTC 287
QY 288 TCTTCACTCTCGCTGAAAGGCTCCAGGCACCACTTCCAGCTGGATCTGAGGACTTCCAG 347
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
288 CTGCACTACCGAATTTTCAGCATTTTGGACAGGAATGCACTTTAGAACTTTAAGCCCTCGCG 347
QY 348 CAGCCTAGTGGCTCTCTGGCTTTATTTGTGACAGCTTGGGAAAGACAGGACACTAAGTCTGT 407
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
348 GA---TTTTGACAGTCACTTTATTTGTCCAGGTACTTGGAAAGATGGTGCT---TCAGA 401
QY 408 GCAGACTTTTACCGCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGCGATTCACACAGAAA 467
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
402 GACTCAGAAACCCGAGGTGCAGAAATGCTTCTATCAGGGAATTTATCAGAAATGCACAGCTC 461
QY 468 CTCCTCAGTGGCCCTTTCAACCTGCCAAGGCTTGTTCAGGCATGATACGACAGAGAGGC 527
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
462 CTCCTCTGTGCTGTGTCTACGTGTGCTGGCTTGTTCAGGTTTAAATGAAGACACGAAAAA 521
```

Qy	528	AGATTACTTCTCTAAGGCCACTTCTCTTCCACACCTCTCATGTGAARACTCGGCAGAGCTGCCCA	587
Db	522	TGAATTCCTCATCTCGCCATTACTCAGCTTTGGCCGAGAAACAACACACAGCTCCCC	591
Qy	588	AGCAGCTCGCCATCCACGTACTGTACAAGAGATCCACAGAGCCCATGCTCTCTGGGCG	647
Db	582	TGCGGGTACCATCTCACGTACTGTACAAAGGACAGCAGAGGAGAAATCCAGCGGTA	641
Qy	648	CAGTAGGCTCTGGTACCTCAAGG-----CATGGAGCTGGCAGATCAACCCCTCGAC	702
Db	642	CCGTGGCTACCCGGCTCTGGCCGGAATTATCTGTGTTACTCCCCAAGTCAATTCGCCA	701
Qy	703	AGCAGC-----GACCTTCGCCCTGGGACTGCCACAAAAGCAGCATTTCTG	746
Db	702	TGCATCTCAGAGTCSGAGACAGAGTATCACCATGGAAGTTGCAAAAGCAGCATTTTGT	761
Qy	747	TGGAAGACCAAGAAATACATGCCCGAGCTCCCAAGAGACCTTCTCATCTTCCGAGA	806
Db	762	TGGACGACGCAAGAAATATGCTCCAGGCTCCCAAGAGCAGGACACTATCTAAGGTTTGA	821
Qy	807	TGAGTATAAGTCTTGCTTACGGCATAAAGCGCTCTCTTCTGAGGTCCCATAGAAATGAAGA	866
Db	822	TGAATATGGGAGCTCTGGGCGACCCAGAAGATCACTCGGAAATACAAA-----AGGG	875
Qy	867	ACTGAAGCTGGAGACCTTGGTGGTGGTGCAGAAAGATGATGCAAAACCATGCGCATGA	926
Db	876	CCTCATGTGGAAACCTCTGGTGGGAGACAAGAAATGGTGGAAAGCATGCGCAAGG	935
Qy	927	AAATATCACCACTTACGTCTCACGATACTCAACATGGTATCTGCTTTATTCAAGATGG	986
Db	936	AAATGTCACCATACATCTCACAGTAATGAACATGGTTTCTGGCCTATTTAAAGATGG	995
Qy	987	AACATATGAGGAAACATCAACATTGCAATTTGTAGTCTGATCTTCTAGAAGATGAACA	1046
Db	996	GACTATTGGAAGTGCATAAAACGTGGTTGTGGTGAGCCTAATCTTCTGGAACAAGAACC	1055
Qy	1047	GCCAGGACTGTGTAGTACCAACGACGACACCACTTAAAGTAGCTTCTCGCAGTGCGCA	1106
Db	1056	TGAGGATTTATGATCAACCATATGACAGACCACTCTCTGAATAGTTTTGTCAATGGCA	1115
Qy	1107	GTCTGGATTGATGGGGAAGATGGGACTCGTATGACACGCCATCTTACTGATGCTCT	1166
Db	1116	GTCTGCCCTCATTTGGAAGAAATGGCAAGACACATGATCATGCCATCTTACTAAGGAT	1175
Qy	1167	GGATATATGTTCTGGGAAGATGAGCCCTGTGACACTTTGGGATTTGCAACCATAGTGG	1226
Db	1176	TGATATTGTTCTTGGGAAGATGAACCATGTGACATCTAGGGTTTGGCCCATCATGTGG	1235
Qy	1227	AATGTGTAGTAATATCGCAGCTGCACGATTAATGAAGATACAGGCTTTGGACTGGCCTT	1286
Db	1236	AATGTGCTCTAGTACCGAAGTTGTACCATCAATAGGACACAGGACTTTGGCCTTGGCCTT	1295
Qy	1287	CACCATTTGCCCATGAGTCTGGACACAACCTTGGCATGATTCATGTGGAGAGGGGAACAT	1346
Db	1296	CACCATCGTCTAGTCTAGGCGCAACCTTTGGTATGATTCAGACGAGAGAGGGAATCC	1355
Qy	1347	GTGTAAAGTCCGAGGGCAACATATGTCCCCATATTGGCAGGACGCAATGGAGTCTT	1406
Db	1356	CTGCAGAAAGGCTGAAGGCAATATCATGTCTCCACACTGACCGGAAACAATGGAGTGT	1415
Qy	1407	CTCCTGGTCACTCGAGCGCCAGTATCTACAAATTTCTAAGCACCGCTCAAGCTAT	1466
Db	1416	TTCATGCTCTTTGAGCGCGCAGTATCTCAAGAAATTTCTCAGCACACCTCAGCGGG	1475
Qy	1467	CTGCCCTTGTGATCGACCAAAAGCCTGTGTAAGGAATACAAGTATCTGTAGAAATTCGCCAGG	1526
Db	1476	GTGTCTAGTGGATGAGCCCAAGCAAGCAGGACAGTATAAATATCCGCAAACTACCAAG	1535
Qy	1527	AGAAATTATATGATGCAACACACAGTGCAGTGGCAGTTTCGAGAGAAGCCAGCTCTG	1586
Db	1536	ACAGATTATATGCTGACACAGTGTAAATGGCAATTTGGAGCAAAAGCCAGGTATG	1595

QY	1587	CATCGTCGACCTTTA	AAAGGACATCTGTAAAGCCCTGTGCTGCCCATCGTATTGGAAAGGAA	1646
Db	1596	CAGCCTTGGTTTTGT	CAAGGATATTTGCAATTCATTTTGGTGGCCACCGAGTAGGCCACAG	1655
QY	1647	ATGTGAGACTAAATTTAT	GCCAGCAGCAAGAGGCACAAATTTGTGGGCATGACATATGTGGTG	1706
Db	1656	GTGTGAGACCAAGTTTAT	CCCGCAGCAAGAGGCACAAATTTGTGGGCATGACATATGTGGTG	1715
QY	1707	CCGGGAGAGCAGTGTGT	GAATAATGGTGATGAAGGCCCAAGCCCAACCCATCGGCACATG	1766
Db	1716	TCGGCAAGCCAGTGCCT	TAAGTTTGGGAGCTCGGGCCCCGGGCCATCCACGGCCAGTG	1775
QY	1767	GTCCGACTGGTCTTCT	TGGTCCCATGCTCCAGGACCTCGGAGGGGAGTATCTCATAG	1826
Db	1776	GTCCGCTGGTCGAAGT	GGTCAGAAATGTTCCCGGACATGTGGTGGAGGAGTCAAGTTC	1835
QY	1827	GAGTCGCTCTGCA	CCCAACCCATCGCATGGAGGGAATTTCTGTGAGGGTCCAC	1886
Db	1836	GGAGAGACACTTGC	ATAAACCAAGCCTCAGTATGGTGGCATATCTGTCCAGGTTCTAG	1895
QY	1887	TGCGACTCTGAAGCT	CTGCCACATGCTCCCGGACAGTGTTGACTTCCTCGTGCC	1946
Db	1896	CCGTATTTATCAGCT	GTGCAATTAACCCCTTCSAATGAATAGCTTGGATTTTCGGGC	1955
QY	1947	TGCTCAGTGTCCG	AGCACACAGCAGATTCAGAGCGCGCATCAAGTGGAAAGCC	2006
Db	1956	CCAACAGTGTG	CAGAAATACAGCAAACTTTCCGTGGATGGTTCTACCACTGGAAACC	2015
QY	2007	TTACACTCAGTAGA	AGATCAGGACTTATGCAAACTCTACTGTATCGCAGAGGATTTGA	2066
Db	2016	CTATACAAAGT	TGGAAGGAAGATCGATGCAAACTGTACTGCAAGGCTGAGAACTTTGA	2075
QY	2067	TTTCTTCTTTTCT	TGTCAAATAAAGTCAAAGATGGGACTCAATCTCGGAGGATAGCCG	2126
Db	2076	ATTTTTTTTTTGC	AATGTTCGGCAAGATGAAGATGGAACTCCCTGCTCCCAACAGAA	2135
QY	2127	TAATGTTGTATAG	TGGATATGTGAGAGAGTTCGATGTGCAATGTCTTTCGATCTCTAA	2195
Db	2136	TGATGTTGTATTGA	CGGGTTTGTGAACAGTGGGATGTGATCATGAACTAGGCTCTAA	2246
QY	2187	TGCTGTTTGAAG	AGCTGTGGGGTGTGTAAACGGGAATTAATCAGCTGCACGATTCACAG	2246
Db	2196	AGCAGTTTCAGAT	GTCTGTGGGGTTGCAAAAGGTGATAATTAACATTTGCAAGTTTTATA	2255
QY	2247	GGGTCCTTACA	CAAGCACCAACACCAACAGATTAATCACAATGTGATCAGTGTCCCTTC	2306
Db	2256	AGGCTGTACTCA	ACACAGCAATAAAGCAATATATTCGGTGTCTATCATCTTCAGC	2315
QY	2307	TGAGGCGCGGAT	TCGGATCTATGAATGAACGTCTCTACCTCTCACAATTTCTGTGGC	2366
Db	2316	TGGGCCCCAGAG	CAATGCAATCCAGGAGCTCGAGTTTCTCTCCAGTTACCTCGAGTTCG	2375
QY	2367	CAATGCCCTCAGA	AGGTACTACTGAAATGGGCACTGGACGTTGGAGCTGCGCCCGCGGTA	2426
Db	2376	AAGCCTCAGTCA	AAAAGTATTACCTCACCGGGGCTGGAGCATGACTGGGCTGGGAGTT	2435
QY	2427	CAAAATTTTCGG	GCACATACTTTTCGACTACAGACGCTCTATATGAGCCGAGAACTTAAT	2486
Db	2436	CCCCTTCGCTGG	AGCACAGCTTTTGAATACACGCGCTCTTCAAACCGCCCGGACGTCGTA	2495
QY	2487	CGCTACTGAG	CAACCAACAGACACTGATTGTGGAGCTGCTGTTTCAGGGAGGAAACC	2546
Db	2496	CGCGCCAGG	CCCCCAAAATGAGACGCTTGGTCTTTGAAATTTCTGATGCAAGGCCAAAATCC	2555
QY	2547	GGGTGTTGCTTGG	GAATACTCCATGCTTCGCTTGGGGAACGGAAGACAGCCCTTCGCCCA	2606
Db	2556	AGGGATAGCTTGG	AAGTATGCACTTCCAAAGGTCATGAATGGAACTCCACCAAGCCACAA	2615
QY	2607	---	GCCCCAGCTACACTTGGGCACTCGTGCCTCTGAGTCTCGTGCTCGGAGGGG	2663
Db	2616	AAGACCTGCCT	TATACCTTGGATATCGTGCAGTCAAGTCTCGTCTCTGTGGTGGAG	2675
QY	2664	ACAGATGACCG	TGAGAGGGCTGTACAGAGACTGAAAGTTTCAAGTAAATATATCTCCTT	2723

Db	2676	TTACATAAATGTAAAGGCCAATTGCTTTGCAGAGTCAAAATACTCAAAGTCAATTCCTCAATT	2735
Qy	2724	CTGCAATCCCAAGACACGACTGTGTCACGGGCTGGTGCCTTGCAAAAGTATCTGCCTGTCC	2783
Db	2736	CTGCAGTGCAAAACCAAGCCAGTAATGAGCCCAAAATCTGCAACGCTTCTCTCTGCC	2795
Qy	2784	TCCCAGCTGTGTCGTGGGGAACTGGAGTGGCTGCAAGTCGACGTGTGTGGCGGGGTGCCCA	2843
Db	2796	GGCTTACTTGATGCGCAGGTGAATGGAGTACATGTAGCAAGGCCCTGTGTCTGGAGGCCAGCA	2855
Qy	2844	GAGCGGCCCGTGCAGTGCACACGGCGGGTGCACT---ATGACTCGGAGCGAGTCCCGGC	2900
Db	2856	GAGCCGAAAGNTCCAGTGTGTGCAAAAGAGCCCTTCCAAAAGGAGGACGATGTTGCA	2915
Qy	2901	CAGCCTGTGCGCTCAGCCTGTCTCCCTCCAGAGCGAGGCTTGCACATCTTCAGAGCTGCCC	2960
Db	2916	TTCTCTGTCTCAGTGCAGCACACCCACTCAGGTCCAGGCTTGCACAGCCATGCCTGTCC	2975
Qy	2961	ACCTGCATGAGCGCGCGGCCCTGGGCAGAGTGTCTACACCTGTGGGAAGGSGTGGAG	3020
Db	2976	TCCCAATGAGACCTTGCACCTGGTCTCAGTGTTCAGAACCTGTGGACGAGGTGAG	3035
Qy	3021	GAAGCGGCAGTGGCCTGTGAAGACACCAACCCCTCGGCCAGAGCGAGCTGTGCCCGA	3080
Db	3036	GAAGCGTGAATCCTCTGTCAAGGCTCTGGC-----GCAGAAACCTCTCCCGA	3083
Qy	3081	CGCTGTCTGCACCTCCGAGCCCAAGCCAGGATGCATGAAGCTGTCTGTCTCAGCGCTG	3140
Db	3084	GAGCCAGTGTACCAGTCTCCCCAGACTGCAGAGCTGCAGGAGGGCTGTGTGCTTGGACGATG	3143
Qy	3141	CCAAAGCCCAAGAGCTGCAGTGGCTGTGTCCGCTGTCTCCAGTGTCTCTGTGACATG	3200
Db	3144	CCCCAAGACAGCCGGCTACAGTGGTCTCTTCTGTGGAGCGAGTGTCTTGCACCTG	3203
Qy	3201	TGAAAGAGGAACACAGAAAAGATCTTAAATATGTCTGTAAGATATGTTCTTGAAAGTA	3260
Db	3204	TGCTTTGGGTGTGAGGAAGAGGAGATGAAATGTCAGCGAGAGGGCTTCCAGGGAAAGCT	3263
Qy	3261	TCGAGAGCTGGCCTCAAAGAAGTGTCTCACATTTGCCGAAGCCAGCTGGAGCTGGAACG	3320
Db	3264	GATAACTTTCCAGAGCGAAGTGCCTGTAATTTAAGAAACCAATCTGACCTTGGNAGA	3323
Qy	3321	TGCTTGCGCCCGCTTCCATGCCCCAGGCAACCCCAATTTGTCTGTGCGGACCCCTCGAG	3380
Db	3324	GACTGTCAACCGACGGCTTGTCCCGAGCCA-----TCCAGTGTACAACATGGT	3371
Qy	3381	GGGCAGCTGTGTTCCTCACCTGGTCTCAGTGCACGGCCAGCTGTGGGGAGGCGCTTCA	3440
Db	3372	AGCTGGATGTGATTCATGTGCGTGGACAGTGCACAGTCACTGTGGGGGAGGGGTCCA	3431
Qy	3441	GAGAGGTCCGTGCAGTGCCTTGGCTGGCGGGGCCCGGCCCTCAGGCTGCCCTCTGCACCA	3500
Db	3432	GACCGGTTCAGTCCACTGTGTTCAGCAAGGCCGGCTTCTCCTCAAGTTGTCTGCTCCATCA	3491
Qy	3501	GAGCCTTGGCCTTCCCTGGCGTGCACACTCACTTCTTGCCCCCAITGTCAGAGAAGA---A	3557
Db	3492	GAAACTCCGGTGTACGAGCCTGTATACAAACTTCTGTCCAGCTCCTGTAAAAGAGAGA	3551
Qy	3558	AGATGCCCTTCTGCAAAAGACTACTTCCACTGGTGTCTACTGGTATCCCGACCCAGCATGGGATGTG	3617
Db	3552	GGATCCATCTCGGTAGATTCTTCACTGGTGTACCTAGTTGCTCTCAGCATGGTGTCTG	3611
Qy	3618	CAGCCACAAGTTCTACGGCAAGCAGTGTCTGCAAGACTTGTCTTAAG	3663
Db	3612	CAACCAAGTTTTTACCGGAAAAAATACTCTCAAGTTCATGACACAAGG	3657

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2005, 16:34:24 ; Search time 15131 Seconds
(without alignments)
11768.748 Million cell updates/sec

Title: US-10-804-457-3
Perfect score: 3675
Sequence: 1 atgaagcccgccgcgcggg.....gctctaagtccaactgtga 3675

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sta.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3675	100.0	3675	6	AR229456
2	3675	100.0	3675	6	AR534556
3	3675	100.0	4042	6	AR229457
4	3675	100.0	4042	6	AR534557
5	3673	100.0	3675	6	BD170564
6	3671	99.9	3675	6	AX1319852
7	3671	99.9	5610	6	BD170080
8	3668	99.8	8435	6	BD170083
9	3668	99.8	8505	6	BD170084
10	3182	86.6	3227	9	HS315734
11	3135	85.3	4888	6	AX786884
12	3064	83.4	4234	6	BD170079
13	3064	83.4	4234	9	AB095949
14	2849	77.5	7668	6	BD170085
15	2692	73.3	3287	6	CQ850584
16	2692	73.3	3287	9	AK127746
17	2544	69.2	4904	10	AK173338
18	2055	55.9	2997	6	AX574590
19	2022	55.0	2902	6	AX574594

20	1874.6	51.0	2895	6	AX574596	AX574596	Sequence
21	1850.2	50.3	3402	10	BC034739	BC034739	Mus muscu
22	1632.8	44.4	1668	6	AX481380	AX481380	Sequence
23	1567.8	42.7	2166	6	AX763026	AX763026	Sequence
24	1460.8	39.7	2225	9	AK122980	AK122980	Homo sapi
25	1449.8	39.5	1476	6	AR229455	AR229455	Sequence
26	1449.8	39.5	1476	6	AR534555	AR534555	Sequence
27	1309.8	35.6	1313	6	CQ738549	CQ738549	Sequence
28	1141	31.0	5130	9	BC063283	BC063283	Homo sapi
29	1139.4	31.0	3666	6	BD170498	BD170498	Novel agg
30	1031.8	28.1	2805	6	AX319854	AX319854	Sequence
31	1010.6	27.5	3389	6	CQ819033	CQ819033	Sequence
32	988.2	26.9	3445	9	HS311903	HS311903	Homo sapi
33	978.2	26.8	2433	6	AX574592	AX574592	Sequence
34	912.2	24.8	4072	6	CQ845833	CQ845833	Sequence
35	912.2	24.8	4072	9	AK131403	AK131403	Homo sapi
36	805.6	21.9	2172	6	AX799384	AX799384	Sequence
37	766	20.8	2247	6	AX763025	AX763025	Sequence
38	417.6	11.4	545	6	BD170082	BD170082	Novel ADA
39	363	9.9	2417	9	AK122677	AK122677	Homo sapi
40	323.8	8.8	82998	2	AC022970	AC022970	Homo sapi
41	323.8	8.8	104776	9	AC022424	AC022424	Homo sapi
42	323.8	8.8	187084	9	AC010269	AC010269	Homo sapi
43	323	8.8	762	6	CQ716585	CQ716585	Sequence
44	320.6	8.7	1528	10	BC046506	BC046506	Mus muscu
45	320.2	8.7	422	6	BD170081	BD170081	Novel ADA

ALIGNMENTS

RESULT 1	AR229456	Sequence 3 from patent US 6448388.	3675 bp	DNA	linear	PAT 20-DEC-2002
LOCUS	AR229456					
DEFINITION	Sequence 3 from patent US 6448388.					
ACCESSION	AR229456					
VERSION	AR229456.1	GI:27268950				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 3675)					
AUTHORS	Fridde, C.J. and Hilbun, E.					
TITLE	Human proteases and polynucleotides encoding the same					
JOURNAL	Patent: US 6448388-A 3 10-SEP-2002;					
FEATURES	Location/Qualifiers					
source	1..3675					
	/organism="unknown"					
	/mol_type="genomic DNA"					

ORIGIN	Query Match	100.0%	Score 3675;	DB 6;	Length 3675;
	Best Local Similarity	100.0%	Pred. No. 0;		
	Matches 3675;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGAAGCCCGCGCGCGGATGGCGGGCTTTGGCGGCGCTGTGGATGCTGTGGCGCAG	60		
Db	1	ATGAAGCCCGCGCGCGGATGGCGGGCTTTGGCGGCGCTGTGGATGCTGTGGCGCAG	60		
Qy	61	GTGGCGGAGCAGCAGCTTGGTGGATGGACCGGCGGCGGCGGCGGCGGCGGCGGCGG	120		
Db	61	GTGGCGGAGCAGCAGCTTGGTGGATGGACCGGCGGCGGCGGCGGCGGCGGCGGCGG	120		
Qy	121	AGCGTCCCGCGCTCTCTCCACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	180		
Db	121	AGCGTCCCGCGCTCTCTCTCCACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	180		
Qy	181	GACCTGGTCTCTGGCTTACGAGTTACCAACAGGCGGCGGCGGCGGCGGCGGCGGCGG	240		
Db	181	GACCTGGTCTCTGGCTTACGAGTTACCAACAGGCGGCGGCGGCGGCGGCGGCGGCGG	240		
Qy	241	CACCATCAGCGCGGAGAGAGCAGTGGCGGCTGTGGAGTTGAGTCTCTTACCTTCGG	300		

Db 241 CACCATCAGGGCGGAGAGACGAGTGCGCGTGTCCGAGGTTGAGTCTCTTCACTTCGG 300
Qy 301 CTGAAGGCTCCAGGCA CGACTTTCACGTGATCTGAGGACTTCAGCAGGCTTAGTGGCT 360
Db 301 CTGAAGGCTCCAGGCA CGACTTTCACGTGATCTGAGGACTTCAGCAGGCTTAGTGGCT 360
Qy 361 CTTGGCTTTATTTGTCAGACGTTGGGAAAGACAGGCACTAAAGTCTGTGTCAGACTTTACCG 420
Db 361 CTTGGCTTTATTTGTCAGACGTTGGGAAAGACAGGCACTAAAGTCTGTGTCAGACTTTACCG 420
Qy 421 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGCGATCACACAGAACTCTCAGTGGCC 480
Db 421 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGCGATCACACAGAACTCTCAGTGGCC 480
Qy 481 CTTTCAACCTGCGAAGGCTTGTCAAGGATGATGAAACAGAAAGGAGGAGATTACTTCCTA 540
Db 481 CTTTCAACCTGCGAAGGCTTGTCAAGGATGATGAAACAGAAAGGAGGAGATTACTTCCTA 540
Qy 541 AGGCCACTTCTTCACACCTCTCATGGAAATCTCGGACAGAGTCCCAAGGACGCTGCCA 600
Db 541 AGGCCACTTCTTCACACCTCTCATGGAACTCGGACAGAGTCCCAAGGACGCTGCCA 600
Qy 601 TCCACGCTACTGTACAGAGATCCACAGAGCCCATGCTCTCGGGCCAGTGAAGTCCCTG 660
Db 601 TCCACGCTACTGTACAGAGATCCACAGAGCCCATGCTCTCGGGCCAGTGAAGTCCCTG 660
Qy 661 GTGACCTCAAGGACATGGGAGCTGGCACATCAACCCCTGACAGCAGGACCTTCGCTG 720
Db 661 GTGACCTCAAGGACATGGGAGCTGGCACATCAACCCCTGACAGCAGGACCTTCGCTG 720
Qy 721 GGACTGCCCAAAAGAGAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCCAGACCTCCC 780
Db 721 GGACTGCCCAAAAGAGAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCCAGACCTCCC 780
Qy 781 AAGGAAGACCTCTTCATCTTGCAGATGAGTAAAGTCTTGCTTACGGGATTAAGGCTCT 840
Db 781 AAGGAAGACCTCTTCATCTTGCAGATGAGTAAAGTCTTGCTTACGGGATTAAGGCTCT 840
Qy 841 CTTCTGAGGTCCTATGAAATAGAACTGAAGCTGGAGACCTTGGTGGTGGTGCACAAA 900
Db 841 CTTCTGAGGTCCTATGAAATAGAACTGAAGCTGGAGACCTTGGTGGTGGTGCACAAA 900
Qy 901 AAGATGATGAAACCAATGCCATGAAATATACACCTAGCTGCTACAGATCTCAAC 960
Db 901 AAGATGATGAAACCAATGCCATGAAATATACACCTAGCTGCTACAGATCTCAAC 960
Qy 961 ATGGTATCTGCTTTATTCAAAGATGAAACATAGAGAGAAACATCAACATTCGAATTTGA 1020
Db 961 ATGGTATCTGCTTTATTCAAAGATGAAACATAGAGAGAAACATCAACATTCGAATTTGA 1020
Qy 1021 GGTCTGATTTCTAGAAATGAAACGAGGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1080
Db 1021 GGTCTGATTTCTAGAAATGAAACGAGGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1080
Qy 1081 ACTTAAAGTAGCTCTGCCAGTGGCAGTCTGGATTTGATGGGAAAGATGGGACTCGTCAT 1140
Db 1081 ACTTAAAGTAGCTCTGCCAGTGGCAGTCTGGATTTGATGGGAAAGATGGGACTCGTCAT 1140
Qy 1141 GACCAAGCCATCTTACTGATCTGGTCTGGATATAAGTCTGGGAAAGATGGGACTCGTCAT 1200
Db 1141 GACCAAGCCATCTTACTGATCTGGTCTGGATATAAGTCTGGGAAAGATGGGACTCGTCAT 1200
Qy 1201 ACTTTGGGATTTGCAACCAATAGTGGATGTGTAGTAAATATCCAGCTGACAGATTAAT 1260
Db 1201 ACTTTGGGATTTGCAACCAATAGTGGATGTGTAGTAAATATCCAGCTGACAGATTAAT 1260
Qy 1261 GAAGTACAGGCTTTGAGTGGCCTTACCAATGGCCATGAGTCTGGACACAACTTTGGC 1320
Db 1261 GAAGTACAGGCTTTGAGTGGCCTTACCAATGGCCATGAGTCTGGACACAACTTTGGC 1320
Qy 1321 ATGATTTCAATGGAAGAGGAAACATGTGTAAGTCCGAGGCAACATCATGTCCCT 1380
Db 1321 ATGATTTCAATGGAAGAGGAAACATGTGTAAGTCCGAGGCAACATCATGTCCCT 1380

Qy 1381 ACATTTGGCAGGACGCAATGGAGTCTTCTCTGTCACCTGCGCGCCAGTATCTTACAC 1440
Db 1381 ACATTTGGCAGGACGCAATGGAGTCTTCTCTGTCACCTGCGCGCCAGTATCTTACAC 1440
Qy 1441 AAATTTCTAAGCAGCCGCTCAAGCTATCTGCTTCTGATCAGCCAAAGCTCTGTAAGAA 1500
Db 1441 AAATTTCTAAGCAGCCGCTCAAGCTATCTGCTTCTGATCAGCCAAAGCTCTGTAAGAA 1500
Qy 1501 TACAAGTATCTTGAGAAATTTGCCAGAGAAATATATGATGCAACACACAGTGCAGTGG 1560
Db 1501 TACAAGTATCTTGAGAAATTTGCCAGAGAAATATATGATGCAACACACAGTGCAGTGG 1560
Qy 1561 CAGTTTCGGAGAGAAAGCCCAAGCTCTGATGCTGGACTTTTAAAGGAATCTGTAAGGCC 1620
Db 1561 CAGTTTCGGAGAGAAAGCCCAAGCTCTGATGCTGGACTTTTAAAGGAATCTGTAAGGCC 1620
Qy 1621 CTGTGCTGCCATCTGATTTGGAAGGAAATGTGAGACTAAATTTATGCCAGCAGCAAGGCC 1680
Db 1621 CTGTGCTGCCATCTGATTTGGAAGGAAATGTGAGACTAAATTTATGCCAGCAGCAAGGCC 1680
Qy 1681 ACAATTTCTGGGACATGACATGTGCTGCGGGAGGACAGTGTGAAATATGTTGATGAA 1740
Db 1681 ACAATTTCTGGGACATGACATGTGCTGCGGGAGGACAGTGTGAAATATGTTGATGAA 1740
Qy 1741 GGGCCAAAGCCCAACCCATGGCCACTGCTCGGACTGCTTCTTGGTCCCAATGCTCCAGG 1800
Db 1741 GGGCCAAAGCCCAACCCATGGCCACTGCTCGGACTGCTTCTTGGTCCCAATGCTCCAGG 1800
Qy 1801 ACCTGCGGAGGGGAGATCTCATAGAGTGCCTCTGCAACCAACCCCAAGCCATCGCAT 1860
Db 1801 ACCTGCGGAGGGGAGATCTCATAGAGTGCCTCTGCAACCAACCCCAAGCCATCGCAT 1860
Qy 1861 GGAGGGAAGTCTGTGAGGGCTCCACCTCGCACTCTGAAGCTCTGCAACAGTCTGAAATGT 1920
Db 1861 GGAGGGAAGTCTGTGAGGGCTCCACCTCGCACTCTGAAGCTCTGCAACAGTCTGAAATGT 1920
Qy 1921 CCCCAGGACAGTGTGACTTTCCTGCTGCTCAGTGTGCCGAGCAACACAGCAGCAGATTC 1980
Db 1921 CCCCAGGACAGTGTGACTTTCCTGCTGCTCAGTGTGCCGAGCAACACAGCAGCAGATTC 1980
Qy 1981 AGAGGCGGCACTACAAAGTGAAGCTTACCTCAAGTGAAGATCAGACTTTATGCAAA 2040
Db 1981 AGAGGCGGCACTACAAAGTGAAGCTTACCTCAAGTGAAGATCAGACTTTATGCAAA 2040
Qy 2041 CTCTACTGTATCGCAGAAAGATTTGATTTCTTCTTTTCTTCTTCTTCTTCTTCTTCTT 2100
Db 2041 CTCTACTGTATCGCAGAAAGATTTGATTTCTTCTTTTCTTCTTCTTCTTCTTCTTCTT 2100
Qy 2101 GGGACTCCATGCTCGGAGGATAGCCGTAATGTTGTTGATAGATGGGATATGTGAGAGTT 2160
Db 2101 GGGACTCCATGCTCGGAGGATAGCCGTAATGTTGTTGATAGATGGGATATGTGAGAGTT 2160
Qy 2161 GGAATGACAAATGCTCTGGATCTGATGCTGTGAAGAGCTCTGTGGGGTGTGAACGG 2220
Db 2161 GGAATGACAAATGCTCTGGATCTGATGCTGTGAAGAGCTCTGTGGGGTGTGAACGG 2220
Qy 2221 AATAACTCAGCCTGCAACGATTCACAGGGGTCTCTACACCAAGCAGCACCACCAACCG 2280
Db 2221 AATAACTCAGCCTGCAACGATTCACAGGGGTCTCTACACCAAGCAGCACCACCAACCG 2280
Qy 2281 TATTTACATGGTCAACCAATCTTCTTCTGAGCCCGGAGTATCCGATCTTATGAATGAAC 2340
Db 2281 TATTTACATGGTCAACCAATCTTCTTCTGAGCCCGGAGTATCCGATCTTATGAATGAAC 2340
Qy 2341 GTCTCTACCTCTACATTTCTGTGGGCAATGCCCTCAGAAAGTACTACTGAAATGGGAC 2400
Db 2341 GTCTCTACCTCTACATTTCTGTGGGCAATGCCCTCAGAAAGTACTACTGAAATGGGAC 2400
Qy 2401 TGGACGCTGGAATGCGCCCGGGGTACAAATTTTTCGGGCACTACTTTTCGACTACAGCGG 2460
Db 2401 TGGACGCTGGAATGCGCCCGGGGTACAAATTTTTCGGGCACTACTTTTCGACTACAGCGG 2460

QY	2461	TCCTATATATGAGCCCGAGAACTTAATCGCTACTTGGAACAAACCAACGAGACACTGATTTGTG	2520
DB	2461	TCCTATATATGAGCCCGAGAACTTAATCGCTACTTGGAACAAACCAACGAGACACTGATTTGTG	2520
QY	2521	GAGCTGCTGTTTTCAGGGAGAGAACCCCGGTGTTGCTCGGGGAATACTCCATGCGCTCGCTTG	2580
DB	2521	GAGCTGCTGTTTTCAGGGAGAGAACCCCGGTGTTGCTCGGGGAATACTCCATGCGCTCGCTTG	2580
QY	2581	GGGACCGAGAAAGCAGCCCTCTGCCAGCCCAAGCTACACTTTGGGCGCATCTGTCGCTCTGAG	2640
DB	2581	GGGACCGAGAAAGCAGCCCTCTGCCAGCCCAAGCTACACTTTGGGCGCATCTGTCGCTCTGAG	2640
QY	2641	TGCTCTCGTGTCTCGGGAGGGGGAACAGATGACCGTGTGAGAGAGGGCTGCTACAGAGACCTG	2700
DB	2641	TGCTCTCGTGTCTCGGGAGGGGGAACAGATGACCGTGTGAGAGAGGGCTGCTACAGAGACCTG	2700
QY	2701	AAGTTTCAAGTAAATATGTCTCTGCAATCCCAAGACACAGACTGTGCGTGGGGAACCTGGAGTGTCTGAGT	2760
DB	2701	AAGTTTCAAGTAAATATGTCTCTCTGCAATCCCAAGACACAGACTGTGCGTGGGGAACCTGGAGTGTCTGAGT	2760
QY	2761	CCTTGC AAAAGTATCTGCTGTCTCTCCAGCTGTGTCGTGGGGAACCTGGAGTGTCTGCAAGT	2820
DB	2761	CCTTGC AAAAGTATCTGCTGTCTCTCCAGCTGTGTCGTGGGGAACCTGGAGTGTCTGCAAGT	2820
QY	2821	CGGACGTGTGGCGGGGTGCCAGAGCGGCCCGTGTGAGTGTGCAACACGGCGGGTGCATAT	2880
DB	2821	CGGACGTGTGGCGGGGTGCCAGAGCGGCCCGTGTGAGTGTGCAACACGGCGGGTGCATAT	2880
QY	2881	GACTCGGAGCCAGTCCCGGCCAGCGCTGTGCGCTCAGCGCTGCTCCCTCCAGCAGCGCAGGCGC	2940
DB	2881	GACTCGGAGCCAGTCCCGGCCAGCGCTGTGCGCTCAGCGCTGCTCCCTCCAGCAGCGCAGGCGC	2940
QY	2941	TGCAACTCTCTCAGAGCTGCCACCTGTGCATGGAGCGCGGCCCGTGGGAGAGTGTCTCACAC	3000
DB	2941	TGCAACTCTCTCAGAGCTGCCACCTGTGCATGGAGCGCGGCCCGTGGGAGAGTGTCTCACAC	3000
QY	3001	ACCTGTGGGAAGGGGTGGAGAACGGCGGCAAGTGGCTGTGAAGACACCAACCCCTCGGCC	3060
DB	3001	ACCTGTGGGAAGGGGTGGAGAACGGCGGCAAGTGGCTGTGAAGACACCAACCCCTCGGCC	3060
QY	3061	AGAGCGCAGCTGTGCGCGACGCTGTGCAACCTCCGAGCCCAAGGCCAAGCCAGATGCAATGAA	3120
DB	3061	AGAGCGCAGCTGTGCGCGACGCTGTGCAACCTCCGAGCCCAAGGCCAAGCCAGATGCAATGAA	3120
QY	3121	GCCTGTCTGCTTCAGCGCTGCCAACAGGAGCTGCAAGTGTGGTGTGTCGCGCTGG	3180
DB	3121	GCCTGTCTGCTTCAGCGCTGCCAACAGGAGCTGCAAGTGTGGTGTGTCGCGCTGG	3180
QY	3181	TCCCAAGTGTCTGTGACATGTGAAGAGGAACACAGAAAAAGATTTCTTAAATGTGCTGAA	3240
DB	3181	TCCCAAGTGTCTGTGACATGTGAAGAGGAACACAGAAAAAGATTTCTTAAATGTGCTGAA	3240
QY	3241	AAGTATGTTTCTGAAAAGTATCGAGAGCTGGCTCAAGAGAGTGTCTCACATTTGCCGGAAG	3300
DB	3241	AAGTATGTTTCTGAAAAGTATCGAGAGCTGGCTCAAGAGAGTGTCTCACATTTGCCGGAAG	3300
QY	3301	CCGAGCTGTGAGCTGTGGAACCTGCTGGCCCGCTTCCATGCTCCCGAGGACCCCGCCATTT	3360
DB	3301	CCGAGCTGTGAGCTGTGGAACCTGCTGGCCCGCTTCCATGCTCCCGAGGACCCCGCCATTT	3360
QY	3361	GCTGCTCGGGACCTCGAGGGGAGTGTGTTTGTCTCACCTCGTGTCTCAGTGTGCAAGGCGC	3420
DB	3361	GCTGCTCGGGACCTCGAGGGGAGTGTGTTTGTCTCACCTCGTGTCTCAGTGTGCAAGGCGC	3420
QY	3421	AGCTGTGGGGAGCGGTTTACAGAGAGTCTGTGAGTGTGCTTGGCTGGGGGCGCGCGCGCC	3480
DB	3421	AGCTGTGGGGAGCGGTTTACAGAGAGTCTGTGAGTGTGCTTGGCTGGGGGCGCGCGCGCC	3480
QY	3481	TCAGGCTGCTCTGTGCAACGAGGCTTCCGCTTCCATGCTCCCGAGGACCCCGCCATTTT	3540
DB	3481	TCAGGCTGCTCTGTGCAACGAGGCTTCCGCTTCCATGCTCCCGAGGACCCCGCCATTTT	3540
QY	3541	CCCAATTGCAGAGAAAGAGATGCTTCTGTGCAAGAGACTACTTCCACTGTGGTGTACTCTGGTA	3600

Db	3541	CCCATTCGAGAGAAGAATGCTTCTGTGCAAGACTACTTCCATCTGGTCTACCTGGTA	3600
Qy	3601	CCCAGACA CGG GATGTG CAGCCAAAGTCTTACGGCAAGCAGTGTCTGCAAGACTTGTCTCT	3660
Db	3601	CCCAGACA CGG GATGTG CAGCCAAAGTCTTACGGCAAGCAGTGTCTGCAAGACTTGTCTCT	3660
Qy	3661	AAGTCCAACTTGTA	3675
Db	3661	AAGTCCAACTTGTA	3675
RESULT 2			
AR534556			
LOCUS			
DEFINITION Sequence 3 from patent US 6734007.			
ACCESSION AR534556			
VERSION AR534556.1 GI:53924881			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
1. .3675			
Location/Qualifiers			
Patent: US 6734007-A 3 11-MAY-2004;			
Human proteases and polynucleotides encoding the same			
Fridde,C.J. and Hilbun,E.			
1 (bases 1 to 3675)			
Unclassified.			
Unknown.			
Unknow.			
/mol_type="genomic DNA"			
/organism="unknown"			
ORIGIN			
Query Match			
Best Local Similarity 100.0%; Score 3675; DB 6; Length 3675;			
Matches 3675; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGAAGCCCCCGCGCGCGGATGCGGGGCTTGGCGCGCTGTGGAGTGTCTGTGGCGCAG	60
Db	1	ATGAAGCCCCCGCGCGCGGATGCGGGGCTTGGCGCGCTGTGGAGTGTCTGTGGCGCAG	60
Qy	61	GTGGCCGAGCAGCACTGTGTCGCCATGGGACCCGACGGCGCGCGCTGGGAGGCCG	120
Db	61	GTGGCCGAGCAGCACCCTGTGTCGCCATGGGACCCGACGGCGCGCGCTGGGAGGCCG	120
Qy	121	AGGCTCCCGCGTCTCTCTCCACCCGCGGAGCGCGCGGCTGGATGGAAAAAGGCGCAATAT	180
Db	121	AGGCTCCCGCGTCTCTCTCCACCCGCGGAGCGCGCGGCTGGATGGAAAAAGGCGCAATAT	180
Qy	181	GACCTGTCTCTGTCTACGAGTTGACCAAGGGGGGATTA	240
Db	181	GACCTGTCTCTGTCTACGAGTTGACCAAGGGGGGATTA	240
Qy	241	CACCATCAGCGCGGAGAGAGCAGTGGCGTGTCCGAGGTTCAGTCTCTTCACTTCGG	300
Db	241	CACCATCAGCGCGGAGAGAGCAGTGGCGTGTCCGAGGTTCAGTCTCTTCACTTCGG	300
Qy	301	CTGAAAGGCTCCAGGCAACGATTCACGTCGATCTGAGGACTTCACAGCAGCTTAGTGGCT	360
Db	301	CTGAAAGGCTCCAGGCAACGATTCACGTCGATCTGAGGACTTCACAGCAGCTTAGTGGCT	360
Qy	361	CCTGGCTTTATGTGCAGACGTTGGGAAAGACAGGCACTAAGTCTGTGCAGACTTTACCG	420
Db	361	CCTGGCTTTATGTGCAGACGTTGGGAAAGACAGGCACTAAGTCTGTGCAGACTTTACCG	420
Qy	421	CCAGAGACTTCTGTTTCTATCAAGGCTCTTTTCCGATCAACAGAAACTCCTCAGTGGCC	480
Db	421	CCAGAGACTTCTGTTTCTATCAAGGCTCTTTTCCGATCAACAGAAACTCCTCAGTGGCC	480
Qy	481	CTTTCAACCTGCCAAGCTTGTACGAGCATGATACGAACAGAGAGGACAGATTACTTCCTA	540
Db	481	CTTTCAACCTGCCAAGCTTGTACGAGCATGATACGAACAGAGAGGACAGATTACTTCCTA	540
Qy	541	AGGCCACTTCTCTTCAACCTTCACTTCATGAAACTCCGGCAGAGCTGCCCAAGGACGCTGCCCA	600

Db	541	AGGCCACTTCTCTTACACCTCTCATGGAACCTCGGCAGAGCTGCCCAAGCGAGCTCGCCA	600
Qy	601	TCCACGCTACTGTACAAGAGATCCACAGAGCCCCCATGCTCCTGGGGCCAGTGAAGTCTCTG	660
Db	601	TCCACGCTACTGTACAAGAGATCCACAGAGCCCCCATGCTCCTGGGGCCAGTGAAGTCTCTG	660
Qy	661	GTGACCTCAAGGACATGGGAGCTGGGCACATCAACCCCTGCACAGCAGCGACCTTCGCGCTG	720
Db	661	GTGACCTCAAGGACATGGGAGCTGGGCACATCAACCCCTGCACAGCAGCGACCTTCGCGCTG	720
Qy	721	GGACTGCCAAMAAAGCAGCATTTCTGTGAAGACGCGAAGAAATACATGCCCGAGCGCTCTCC	780
Db	721	GGACTGCCAAMAAAGCAGCATTTCTGTGAAGACGCGAAGAAATACATGCCCGAGCGCTCTCC	780
Qy	781	AAGGAAGACTCTTCAATCTTGCAGATGATATAAGTCTTGTCTTACGGCATAAAGCGCTCT	840
Db	781	AAGGAAGACTCTTCAATCTTGCAGATGATATAAGTCTTGTCTTACGGCATAAAGCGCTCT	840
Qy	841	CTTCTGAGTCCCATAGAAATCAAGAACTGAACGTGGAGACCTTGGTGGTGGTGCACAAA	900
Db	841	CTTCTGAGTCCCATAGAAATGAAGACATGAACGTGGAGACCTTGGTGGTGGTGCACAAA	900
Qy	901	AAGATGATGCAAAAACATGGGCCATGAAATATCACCACTACGTGCTCACGATACTCAAC	960
Db	901	AAGATGATGCAAAAACATGGGCCATGAAATATCACCACTACGTGCTCACGATACTCAAC	960
Qy	961	ATGGTATCTGCTTTATTCAAAGATGGAACATAGGAGAAACATCAACATTTGCAATTTGTA	1020
Db	961	ATGGTATCTGCTTTATTCAAAGATGGAACATAGGAGAAACATCAACATTTGCAATTTGTA	1020
Qy	1021	GGTCTGATCTTCTAGAAAGATGAACAGCAGGACTGTGTGATAGTCACCAAGCAGACAC	1080
Db	1021	GGTCTGATCTTCTAGAAAGATGAACAGCAGGACTGTGTGATAGTCACCAAGCAGACAC	1080
Qy	1081	ACCTTAAAGTAGCTTCTGCCAGTGGCAGTCTGGATGTATGGGAAAGATGGGACTGTGTCAT	1140
Db	1081	ACCTTAAAGTAGCTTCTGCCAGTGGCAGTCTGGATGTATGGGAAAGATGGGACTGTGTCAT	1140
Qy	1141	GACACGCCATCTTACGTACTGTCTGGATATATGTTCTTGGAAATAGCCCTGTGAC	1200
Db	1141	GACACGCCATCTTACGTACTGTCTGGATATATGTTCTTGGAAATAGCCCTGTGAC	1200
Qy	1201	ACTTTGGATTTGACACCATAAAGTGAATGTGTAGTAAATATCGCAGCTGCGATTAAT	1260
Db	1201	ACTTTGGATTTGACACCATAAAGTGAATGTGTAGTAAATATCGCAGCTGCGATTAAT	1260
Qy	1261	GAAGATACAGTCTTGGA CTGGCCTTCAACATTTGCCCATGAGTCTGGACACAACTTTGGC	1320
Db	1261	GAAGATACAGTCTTTGGACTTGGCCTTCAACATTTGCCCATGAGTCTGGACACAACTTTGGC	1320
Qy	1321	ATGATTCATGATGAGAAGGGAACATGTGTAAAGTCCGAGGGGAAACATCATGTCCCT	1380
Db	1321	ATGATTCATGATGAGAAGGGAACATGTGTAAAGTCCGAGGGGAAACATCATGTCCCT	1380
Qy	1381	ACATTTGCCAGGACGCAATGGAGTCTTCTCTGTGTCAACCTGCGAGCGCGAGTATCTACAC	1440
Db	1381	ACATTTGCCAGGACGCAATGGAGTCTTCTCTGTGTCAACCTGCGAGCGCGAGTATCTACAC	1440
Qy	1441	AAATTTCTAAGCACCGCTCAAGCTATCTGCTTGTGATCAGCCAAAGCCTGTGAAGAA	1500
Db	1441	AAATTTCTAAGCACCGCTCAAGCTATCTGCTTGTGATCAGCCAAAGCCTGTGAAGAA	1500
Qy	1501	TACAAGTATCCTGAGAAATTTGCCAGGAGAAATATATGATGCAACACACAGTGCAGTGG	1560
Db	1501	TACAAGTATCCTGAGAAATTTGCCAGGAGAAATATATGATGCAACACACAGTGCAGTGG	1560
Qy	1561	CAGTTCCGAGAGAAAGCCAGCTCTCATGCTGGACTTTAAAAGGACATCTGTAAAGCC	1620
Db	1561	CAGTTCCGAGAGAAAGCCAGCTCTCATGCTGGACTTTAAAAGGACATCTGTAAAGCC	1620
Qy	1621	CTGTGTGCCATCGTATTGGAAAGAAATGTGAGACTTAAATTTATGCGCAGCAGCAAGGC	1680
Db	1621	CTGTGTGCCATCGTATTGGAAAGAAATGTGAGACTTAAATTTATGCGCAGCAGCAAGGC	1680

Db	1621	CTGTGGTGGCCATCGTATTGTTGGAAGGAAATGTGAGACTAAATTTATGCGAGCAGCAGAAAGCC	1681
Qy	1681	ACAAATTTGTGGGCATGACATGTGGTGCCGGGGAGGACAGTGTGTGAATATATGTCGATGAA	1740
Db	1681	ACAAATTTGTGGGCATGACATGTGGTGCCGGGGAGGACAGTGTGTGAATATATGTCGATGAA	1740
Qy	1741	GGCCCCAAGCCCAACCATGGGCACTGGTCCGACTGCTCTTTGGTCCCATGCTCCAGG	1800
Db	1741	GGCCCCAAGCCCAACCATGGGCACTGGTCCGACTGCTCTTTGGTCCCATGCTCCAGG	1800
Qy	1801	ACCTGCGGAGGGGAGTATCTCATAGGAGTGCCTCTCTGCACCAACCCCAAGCCATCGCAT	1860
Db	1801	ACCTGCGGAGGGGAGTATCTCATAGGAGTGCCTCTCTGCACCAACCCCAAGCCATCGCAT	1860
Qy	1861	GGAGGGAAGTCTGTGTGAGGGCTCCACTCGCATCTCTGAAGCTCTGTCAACAGTCAGAAATGT	1920
Db	1861	GGAGGGAAGTCTGTGTGAGGGCTCCACTCGCATCTCTGAAGCTCTGTCAACAGTCAGAAATGT	1920
Qy	1921	CCCCGGGACAGTGTGACTTCCGTGCTGCTCAGTGTGCCGAGCACCAACGACGATTC	1980
Db	1921	CCCCGGGACAGTGTGACTTCCGTGCTGCTCAGTGTGCCGAGCACCAACGACGATTC	1980
Qy	1981	AGAGGGCGGCATCAAGTGGAGGCCCTTACACTCAAGTAGAAGATCAGACTTATTCGAA	2040
Db	1981	AGAGGGCGGCATCAAGTGGAGGCCCTTACACTCAAGTAGAAGATCAGACTTATTCGAA	2040
Qy	2041	CTCTACTGTATCGCAGAAGATTGATTTCTTTCTTTTGTGCAATAAAGTCAAGAT	2100
Db	2041	CTCTACTGTATCGCAGAAGATTGATTTCTTTCTTTTGTGCAATAAAGTCAAGAT	2100
Qy	2101	GGGACTCCATGCTCGGAGGATAGCCGTAATGTTGTATAGATGGGATATGTCAGAGCTT	2160
Db	2101	GGGACTCCATGCTCGGAGGATAGCCGTAATGTTGTATAGATGGGATATGTCAGAGCTT	2160
Qy	2161	GGATGTGACAAATGTCCTTGGATCTGATGCTGTTTGAAGACGTCTGTGGGGTGTGTAACGGG	2220
Db	2161	GGATGTGACAAATGTCCTTGGATCTGATGCTGTTTGAAGACGTCTGTGGGGTGTGTAACGGG	2220
Qy	2221	AATAACTCAGCTGTCACGATTTACAGGGGTCTCTACCAAGCACCAACCAACCAACAG	2280
Db	2221	AATAACTCAGCTGTCACGATTTACAGGGGTCTCTACCAAGCACCAACCAACCAACAG	2280
Qy	2281	TATTATCAGATGTCACCAATTCCTTCTGAGCCGGAGTATCCGCATCTATGAATGAAC	2340
Db	2281	TATTATCAGATGTCACCAATTCCTTCTGAGCCGGAGTATCCGCATCTATGAATGAAC	2340
Qy	2341	GTCTCTACTCTTACATTTCTGTGCGCAATGCCCTCAGAAAGTACTACTGAAATGGGCAC	2400
Db	2341	GTCTCTACTCTTACATTTCTGTGCGCAATGCCCTCAGAAAGTACTACTGAAATGGGCAC	2400
Qy	2401	TGGACCTGGAATGGCCCGCCGGTACAAATTTTTCGGGCACCTACTTTCGACTACAGACGG	2460
Db	2401	TGGACCTGGAATGGCCCGCCGGTACAAATTTTTCGGGCACCTACTTTCGACTACAGACGG	2460
Qy	2461	TCCTATAATGAGCCCGAGAACTTAATCGTACTTGACCAACCAACGAGACACTGATTGTG	2520
Db	2461	TCCTATAATGAGCCCGAGAACTTAATCGTACTTGACCAACCAACGAGACACTGATTGTG	2520
Qy	2521	GAGCTGCTGTTTCAGGGAAAGGAACCCGGGTGTGCTCGGAAATATCTCATGSCCTCGCTTG	2580
Db	2521	GAGCTGCTGTTTCAGGGAAAGGAACCCGGGTGTGCTCGGAAATATCTCATGSCCTCGCTTG	2580
Qy	2581	GGGACCGAGAGCAGCCCTCCAGCCCAAGCTACACTTGGGCGATCGTGCCTCTGAG	2640
Db	2581	GGGACCGAGAGCAGCCCTCCAGCCCAAGCTACACTTGGGCGATCGTGCCTCTGAG	2640
Qy	2641	TGCTCCGTGCTCTGCGGAGGGGACAGATGACCGTGAGAGAGGCTCTCTACAGACCTG	2700
Db	2641	TGCTCCGTGCTCTGCGGAGGGGACAGATGACCGTGAGAGAGGCTCTCTACAGACCTG	2700
Qy	2701	AGTTTTCAGTAAATATGTCTTCTTGCAATCCCAAGACAAGACTGTGCACGGGCTGGTG	2760
Db	2701	AGTTTTCAGTAAATATGTCTTCTTGCAATCCCAAGACAAGACTGTGCACGGGCTGGTG	2760

QY	2761	CTTGCAAAATATCTGCTGCTCTCCAGCTGGTCCGTGGGGAACGTGGAGTGCCTGCAAT	2820
Db	2761		
QY	2821	CGGACGTGTGGCGGGGTCCACAGACCGCCCGTGCAGTGCACCGCGGTGCACTAT	2880
Db	2821		
QY	2881	GACTCGAGCCAGTCCCGGCCAGCTGTGCTCCCTCAGCTGCTCCCTCCAGCAGGCGAGCC	2940
Db	2881		
QY	2941	TGCAACTCTCAGAGTCCCACTGTGATGAGGCGCGGCGCTTGGGCAAGTGTCTCAC	3000
Db	2941		
QY	3001	ACCTGTGGGAAGGGGTGGAGGAAGCGGGCAGTGGCTGTAAAGGCAACCAACCCCTCGGCC	3060
Db	3001		
QY	3061	AGAGCGACGTGTGCTGCCAGCGTGTCTGCACTCCGAGCCCAAGCCCAAGGATGCATGAA	3120
Db	3061		
QY	3121	GCCTGTCTGCTTACGCGTGGCCACAAAGCCCAAGAGCTCAGTGTGTCTCCGCTGG	3180
Db	3121		
QY	3181	TCCAGTGTCTGTGACATGTCAAGAGGAAACACAGAAAGATTTCTAAATGTCTGAA	3240
Db	3181		
QY	3241	AAATATGTTCTCGAAAGTATCAGAGCTGGCTCAAGAAAGTGTCTCAATTTGCGAAG	3300
Db	3241		
QY	3301	CCAGCCTGGAGCTGGAACGTGCTGGCGCCCGCTTCCATGCCCCCAGGCAACCCCAATTT	3360
Db	3301		
QY	3361	GCCTGTGGGAGCCCTCGAGGGGAGCTGGTTGGCTTCCATGCCCCCAGGCAACCCCAATTT	3420
Db	3361		
QY	3421	AGCTGTGGGGAGGCGTTTCAAGACGAGTCCGTGCAAGTCCCTGGCTGGGCGCGCGGCC	3480
Db	3421		
QY	3481	TCAGGCTGCTCTCTGCACAGAAAGCTTCCGGCTCCCTGGCTCCCTGGCTGCAACTCTCTGC	3540
Db	3481		
QY	3541	CCATTTGCAGAGAAAGATGCTTCCGAAAGTCTTCCATGCTGCTGCTGCTGCTGCTGCTG	3600
Db	3541		
QY	3601	CCCGAGCAGGATGTGACGACCAAGTCTTACGCAAGCAGTGTGCAAGACTTGTCT	3660
Db	3601		
QY	3661	AACTCCAACTTGTGA	3675
Db	3661		

RESULT 3
AR229457
LOCUS AR229457 4042 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 5 from patent US 6448388.
ACCESSION AR229457
VERSION AR229457.1 GI:27268951
KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4042)
AUTHORS Friddle, C.J. and Hilbun, E.
TITLE Human proteases and polynucleotides encoding the same
JOURNAL Patent: US 6448388-A 5 10-SEP-2002;
FEATURES Location/Qualifiers
source 1..4042
/organism="unknown"
/mol_type="genomic DNA"

Query Match 100.0%; Score 3675; DB 6; Length 4042;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3675; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAAGCCCGCGCGCGGATGGCGGGCTTTGGCGGCGCTGTGGATGCTGTGGCGCAG	60
Db	99	ATGAAGCCCGCGCGCGGATGGCGGGCTTTGGCGGCGCTGTGGATGCTGTGGCGCAG	158
QY	61	GTGGCGGAGCAGGCACTGCGTGGCCATGGGACCCGCGAGCGGCGCTGGGAGCCCG	120
Db	159	GTGGCGGAGCAGGCACTGCGTGGCCATGGGACCCGCGAGCGGCGCTGGGAGCCCG	218
QY	121	AGCGTCCCGCGTCTCTCCACCCCGGAGCGCGGCTGGATGGAAAGGCGCAATAT	180
Db	219	AGCGTCCCGCGTCTCTCCACCCCGGAGCGCGGCTGGATGGAAAGGCGCAATAT	278
QY	181	GACCTGGTCTTGGCTTACGAGTTGACCAAGGGGGGATAGTGTCCCATGAATCATG	240
Db	279	GACCTGGTCTTGGCTTACGAGTTGACCAAGGGGGGATAGTGTCCCATGAATCATG	338
QY	241	CACCATCAGCGCGGAGAGAGCAGTGGCGGTGTCGAGGTTGAGTCTCTTACCTTCGG	300
Db	339	CACCATCAGCGCGGAGAGAGCAGTGGCGGTGTCGAGGTTGAGTCTCTTACCTTCGG	398
QY	301	CTGAAAGGCTCCAGGCAAGCTTCCAGTGGATCTGAGGACTTCCAGAGCCCTAGTGGCT	360
Db	399	CTGAAAGGCTCCAGGCAAGCTTCCAGTGGATCTGAGGACTTCCAGAGCCCTAGTGGCT	458
QY	361	CTGCGCTTATTTGTCAGACGTTGGGAAAGACAGGCACTAAGTCTGTGAGACTTTACCG	420
Db	459	CTGCGCTTATTTGTCAGACGTTGGGAAAGACAGGCACTAAGTCTGTGAGACTTTACCG	518
QY	421	CCAGAGGACTTCTGTTTCTATCAAGCTCTTTGCGATCACACAGAAACTCTCAGTGGCC	480
Db	519	CCAGAGGACTTCTGTTTCTATCAAGCTCTTTGCGATCACACAGAAACTCTCAGTGGCC	578
QY	481	CTTTCAGCTGCAAGGCTTGTTCAGGCTATGATACGAAACAGAGAGGAGAGTACTTCTTA	540
Db	579	CTTTCAGCTGCAAGGCTTGTTCAGGCTATGATACGAAACAGAGAGGAGAGTACTTCTTA	638
QY	541	AGGCCACTTCTTTCACACTCTCATGGAACCTGGGAGAGCTGCCAAGGAGAGTGCACA	600
Db	639	AGGCCACTTCTTTCACACTCTCATGGAACCTGGGAGAGCTGCCAAGGAGAGTGCACA	698
QY	601	TCCACGACTACTGTACAAGAGATCCACAGAGCCCATGCTCTCTGGGCGCAGTGGCTG	660
Db	699	TCCACGACTACTGTACAAGAGATCCACAGAGCCCATGCTCTCTGGGCGCAGTGGCTG	758
QY	661	GTGACCTCAGGACATGGGAGCTGGCAGATCAACCCCTGCAAGAGGAGCCTTCGCTG	720
Db	759	GTGACCTCAGGACATGGGAGCTGGCAGATCAACCCCTGCAAGAGGAGCCTTCGCTG	818
QY	721	GGACTGCCCAAAAGCAGCAATTTCTGTGGAAGACGCAAGAAATATATGCCCCAGGCTCC	780
Db	819	GGACTGCCCAAAAGCAGCAATTTCTGTGGAAGACGCAAGAAATATATGCCCCAGGCTCC	878
QY	781	AGGAGAGACCTTCTCATCTTCCAGATCAGTATAGTCTTGTGCTTACGCGCATAGGCTCT	840
Db	879	AGGAGAGACCTTCTCATCTTCCAGATCAGTATAGTCTTGTGCTTACGCGCATAGGCTCT	938

QY	841	CTTCTGAGGTTCCCATAGAAAATGAGAACTGAACGTGGAGACCTTGGTGGTGGTGGACAAA	900
Db	939	CTTCTGAGGTTCCCATAGAAAATGAGAACTGAACGTGGAGACCTTGGTGGTGGTGGACAAA	998
QY	901	AAGATGATGCAAAACCAATGGCCATGAAAATATATCACCCTAGCTGCTCATCGATATCTCAAC	960
Db	999	AAGATGATGCAAAACCAATGGCCATGAAAATATATCACCCTAGCTGCTCATCGATATCTCAAC	1058
QY	961	ATGGTATCTGCTTTTATTCAAAGATGGAAACAATAGGAGGAAAAATCAACAATTCGCAATTTGTA	1020
Db	1059	ATGGTATCTGCTTTTATTCAAAGATGGAAACAATAGGAGGAAAAATCAACAATTCGCAATTTGTA	1118
QY	1021	GGTCTGATTTCTTCTAGAAAGATGAACAGCCAGGACTGGTGATTAAGTCAACAGCAGACCAAC	1080
Db	1119	GGTCTGATTTCTTCTAGAAAGATGAACAGCCAGGACTGGTGATTAAGTCAACAGCAGACCAAC	1178
QY	1081	ACCTTAAGTAGTCTTCTGCCAGTGGCAGTCTGGATTTGATGGGGGAAGATGGGACTCTGTCAT	1140
Db	1179	ACCTTAAGTAGTCTTCTGCCAGTGGCAGTCTGGATTTGATGGGGGAAGATGGGACTCTGTCAT	1238
QY	1141	GACCAAGCCATCTTATCTGACTGGTCTGGAAATATATGTTCTCTGGAAAGATGAGCCCTGTGAC	1200
Db	1239	GACCAAGCCATCTTATCTGACTGGTCTGGAAATATATGTTCTCTGGAAAGATGAGCCCTGTGAC	1298
QY	1201	ACTTTGGGATTTCCACCCCAATAGTGGAAATGCTGACTGTAATAATCCGACTGCACGATTAAT	1260
Db	1299	ACTTTGGGATTTCCACCCCAATAGTGGAAATGCTGACTGTAATAATCCGACTGCACGATTAAT	1358
QY	1261	GAAATACAGGCTTTGGACTGGCCTTCCACATTCGATGCCATGAGTCTGGACACAACTTTGGC	1320
Db	1359	GAAATACAGGCTTTGGACTGGCCTTCCACATTCGATGCCATGAGTCTGGACACAACTTTGGC	1418
QY	1321	ATGATTTCAATGAGGAGGGAACATGTGTAAGAATCCGAGGGAACATCATGTCCCTT	1380
Db	1419	ATGATTTCAATGAGGAGGGAACATGTGTAAGAATCCGAGGGAACATCATGTCCCTT	1478
QY	1381	ACATTTGGCAGGACGCAATGAGTCTTCTCTGGTCAACCTTCGACGCCCAAGTATCTACAC	1440
Db	1479	ACATTTGGCAGGACGCAATGAGTCTTCTCTGGTCAACCTTCGACGCCCAAGTATCTACAC	1538
QY	1441	AAATTTCTTAAGCAACCGCTCAAGCTATCTGCTTCTGATCAGCCAAAGCTGTGGAAGAA	1500
Db	1539	AAATTTCTTAAGCAACCGCTCAAGCTATCTGCTTCTGATCAGCCAAAGCTGTGGAAGAA	1598
QY	1501	TACAAGTATCTTGAGAAATTCGACGAGAAATATATGATGCAAAACACACAGTGCAGTGG	1560
Db	1599	TACAAGTATCTTGAGAAATTCGACGAGAAATATATGATGCAAAACACACAGTGCAGTGG	1658
QY	1561	CAGTTTCGGAGAGAAAGCAAGCTCTGCATCTGCTGAACTTTAAAAGGACATCTGTAAAGCC	1620
Db	1659	CAGTTTCGGAGAGAAAGCAAGCTCTGCATCTGCTGAACTTTAAAAGGACATCTGTAAAGCC	1718
QY	1621	CTGTGGTGCATCTGATTTGGAAGGAAATGTGAGACTTAAATTTATGCCACGACGACGAAGGC	1680
Db	1719	CTGTGGTGCATCTGATTTGGAAGGAAATGTGAGACTTAAATTTATGCCAGCAGCAGAGGC	1778
QY	1681	ACAATTTCTGGGCAATGACATGTGTGTCGGGAGGAGCATGTGTGAAATATGGTGATGAA	1740
Db	1779	ACAATTTCTGGGCAATGACATGTGTGTCGGGAGGAGCATGTGTGAAATATGGTGATGAA	1838
QY	1741	GGCCCCAAGCCCAACCCATGGCCACTGGTTCGGAATGCTGCTTCTTGGTCCCAATGCTCCAGG	1800
Db	1839	GGCCCCAAGCCCAACCCATGGCCACTGGTTCGGAATGCTGCTTCTTGGTCCCAATGCTCCAGG	1898
QY	1801	ACCTCGGAGGGGAGATCTCATAGAGTGGCTCTGCAACCAACCCCAAGCCATCGCAT	1860
Db	1899	ACCTCGGAGGGGAGATCTCATAGAGTGGCTCTGCAACCAACCCCAAGCCATCGCAT	1958
QY	1861	GGAAGGAGTTCTGTGAGGGCTCACTCGCACTCTGAACTCTGCAACAGTCAGAAATGT	1920
Db	1959	GGAAGGAGTTCTGTGAGGGCTCACTCGCACTCTGAACTCTGCAACAGTCAGAAATGT	2018
. QY	1921	CCCCGGACAGTGTGATCTTCCTGGTGTCTAGTGTGCGGACCAACAGCAGACGATTC	1980

Db	2019		CCCCGGGACAGTGTGTGACTTCGCTCGTCACTAGTGTGCCGAGCACACGACAGACGATTC	2078
Qy	1981		AGAGGCGGCACCTACAAAGTGGAAAGCCTTACACTCAAGTGAAGATCAGGACTTATGCAAA	2040
Db	2079		AGAGGCGGGCACTACAAAGTGGAAAGCCTTACACTCAAGTGAAGATCAGGACTTATGCAAA	2138
Qy	2041		CTCTACTGTATCGCAGAAAGGATTGTGATTTCTTTCTTTCTTTGTCAAAATAAAGTCAAAAGAT	2100
Db	2139		CTCTACTGTATCGCAGAAAGGATTGTGATTTCTTTCTTTGTCAAAATAAAGTCAAAAGAT	2198
Qy	2101		GGGACTCCATGCTCCGGAGGATAGCCGTAATGTTTGTATAGATGGGATATGTGAGAGATTT	2160
Db	2199		GGGACTCCATGCTCCGGAGGATAGCCGTAATGTTTGTATAGATGGGATATGTGAGAGATTT	2258
Qy	2161		GGATGTGACAAATGTCTTTGGATCTGATGCTGTGTGAAGACGTCGTGGGGTGTGTAACGGG	2220
Db	2259		GGATGTGACAAATGTCTTTGGATCTGATGCTGTGTGAAGACGTCGTGGGGTGTGTAACGGG	2318
Qy	2221		AATAACTCAGCTCAGCAAGATTTCACAGGGGTCTCTACACCAAGCACCAACCAACCAAG	2280
Db	2319		AATAACTCAGCTCAGCAAGATTTCACAGGGGTCTCTACACCAAGCACCAACCAACCAAG	2378
Qy	2281		TATTATCACATGGTCAACCATCTCTTCGGAGCCCGAGATCCGCACTCTATGAATGAAC	2340
Db	2379		TATTATCACATGGTCAACCATCTCTTCGGAGCCCGAGATTCGCACTCTATGAATGAAC	2438
Qy	2341		GTCTCTACCTCCTACATTTCTGTGCGCAATGCCCTCAGAAAGTATCTACTGTAATGGCAC	2400
Db	2439		GTCTCTACCTCCTACATTTCTGTGCGCAATGCCCTCAGAAAGTATCTACTGTAATGGCAC	2498
Qy	2401		TGGACCGTGGACTGGCCCGGCCGGTACAAATTTTCGGGCACACTTTTCGACTACAGACGG	2460
Db	2499		TGGACCGTGGACTGGCCCGGCCGGTACAAATTTTCGGGCACACTTTTCGACTACAGACGG	2558
Qy	2461		TCCTTAATAGCCCGGAGAACTTAATCGCTATCTGGACCAACCAACGAGACACTGATTTGT	2520
Db	2559		TCCTTAATAGCCCGGAGAACTTAATCGCTATCTGGACCAACCAACGAGACACTGATTTGT	2618
Qy	2521		GAGCTGTCTTTTCAGGGAAGGAAACCCGGGTGTTCCTGGGAATCTCCATGCTTCGCTTG	2580
Db	2619		GAGCTGTCTTTTCAGGGAAGGAAACCCGGGTGTTCCTGGGAATCTCCATGCTTCGCTTG	2678
Qy	2581		GGGACCGAGACAGCCCTCTGCCACGCCAGTACATTTGGCCCATCTGTCGCTCTGAG	2640
Db	2679		GGGACCGAGAGACAGCCCTCTGCCACGCCAGTACATTTGGCCCATCTGTCGCTCTGAG	2738
Qy	2641		TGCTCCGTGTCTTCGGAGGGGGAGACAGTACCGTGTGAGAGGGGTGCTACAGAGACCTG	2700
Db	2739		TGCTCCGTGTCTTCGGAGGGGGAGACAGTACCGTGTGAGAGGGGTGCTACAGAGACCTG	2798
Qy	2701		AAGTTTCAAGTAATAATGTCTTCTGCAATCCCAAGACACGACCTGTCA CGGGGTCTGTG	2760
Db	2799		AAGTTTCAAGTAATAATGTCTTCTGCAATCCCAAGACACGACCTGTCA CGGGGTCTGTG	2858
Qy	2761		CCTTGCAAAAGTATCTCCCTGTCTCCAGCTGGTCCGTGGGAACTGAGTGCCTGAGT	2820
Db	2859		CCTTGCAAAAGTATCTCCCTGTCTCCAGCTGGTCCGTGGGAACTGAGTGCCTGAGT	2918
Qy	2821		CGGACGTGTGGCGGGGTGCCCAGAGCCGCCCGGTGACGTGCACACGCGGGGTGCATCTAT	2880
Db	2919		CGGACGTGTGGCGGGGTGCCCAGAGCCGCCCGGTGACGTGCACACGCGGGGTGCATCTAT	2978
Qy	2881		GACTCGGAGCAGTCCCGCCAGCTGTGCCCTCAGCTGTCTCCCTCCAGCAGGCGAGGCC	2940
Db	2979		GACTCGGAGCAGTCCCGCCAGCTGTGCCCTCAGCTGTCTCCCTCCAGCAGGCGAGGCC	3038
Qy	2941		TGCACTCTTCAGAGCTGCCCACTTGCAATGGAGCCCGGGCCCTGGGAGAGTGTCTCACAC	3000
Db	3039		TGCACTCTTCAGAGCTGCCCACTTGCAATGGAGCCCGGGCCCTGGGAGAGTGTCTCACAC	3098
Qy	3001		ACCTGTGGGAAGGGTGGAGGAAGGGGACGTGGCCCTGTGAAGGACCAACCCCTGGGCC	3060

```
Db 3099 ACCTGTGGGAGGGGTGGAGGAAGGGGCAAGTGGCTGTAAAGACCAACCCCTCGGCC 3158
Qy 3061 AGAGCCAGCTGTGCCCCGACGTGTCTGCACTCCGAGCCCAAGCCCAAGGATGATGAA 3120
Db 3159 AGAGCCAGCTGTGCCCCGACGTGTCTGCACTCCGAGCCCAAGCCCAAGGATGATGAA 3218
Qy 3121 GCCTGTCTGCTTCAGCGCTGCCACAAGCCCAAGAAAGCTGCAAGTGGCTGTGCTGG 3180
Db 3219 GCCTGTCTGCTTCAGCGCTGCCACAAGCCCAAGAAAGCTGCAAGTGGCTGTGCTGG 3278
Qy 3181 TCCAGTGTCTGTGACATGTGAAAGAGGAACACAGAAAGATTCCTTAAATGTGCTGAA 3240
Db 3279 TCCAGTGTCTGTGACATGTGAAAGAGGAACACAGAAAGATTCCTTAAATGTGCTGAA 3338
Qy 3241 AAGTATGTTTCTGAAAGTATCGAGAGCTGGCCCTCAAGAAAGTCTCACATTTGCCGAAG 3300
Db 3339 AAGTATGTTTCTGAAAGTATCGAGAGCTGGCCCTCAAGAAAGTCTCACATTTGCCGAAG 3398
Qy 3301 CCCAGCTGGAGCTGGAACGTGCTGCGGCCCGCCGCTTCCATGCCCCAGGCAACCCCAATTT 3360
Db 3399 CCAGGCTGGAGCTGGAACGTGCTGCGGCCCGCCGCTTCCATGCCCCAGGCAACCCCAATTT 3458
Qy 3361 GCTGTGCGGACCTCGAGGGGCGAGCTGTTTGGCTTCCCTGCTCAGTGCACGGCC 3420
Db 3459 GCTGTGCGGACCTCGAGGGGCGAGCTGTTTGGCTTCCCTGCTCAGTGCACGGCC 3518
Qy 3421 AGCTGTGGGGAGGCGCTTCAGACAGGTCGCTGCAAGTGGCTGGCTGGGGCGCGCGGCC 3480
Db 3519 AGCTGTGGGGAGGCGCTTCAGACAGGTCGCTGCAAGTGGCTGGCTGGGGCGCGCGGCC 3578
Qy 3481 TCAGGCTGCTCTGCAACAGAGCTTGGGCTTCCCTGCTGCAACTCACTTCTGCG 3540
Db 3579 TCAGGCTGCTCTGCAACAGAGCTTGGGCTTCCCTGCTGCAACTCACTTCTGCG 3638
Qy 3541 CCATTTCAGAGAAAGATGCTTCTGCAAGAGCTACTTCCACTGGTGTACCTGCTA 3600
Db 3639 CCATTTCAGAGAAAGATGCTTCTGCAAGAGCTACTTCCACTGGTGTACCTGCTA 3698
Qy 3601 CCCAGCAGGGATGTGCAGCCACAAAGTTCTACGGCAAGCAGTGTGCAAGACTTGCCTCT 3660
Db 3699 CCCAGCAGGGATGTGCAGCCACAAAGTTCTACGGCAAGCAGTGTGCAAGACTTGCCTCT 3758
Qy 3661 AAGTCAACTTTGGA 3675
Db 3759 AAGTCAACTTTGGA 3773

RESULT 4
AR534557
LOCUS AR534557 4042 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 5 from patent US 6734007.
ACCESSION AR534557
VERSION AR534557.1 GI:53924882
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4042)
AUTHORS Fridde,C.J. and Hilbun,B.
TITLE Human proteases and polynucleotides encoding the same
JOURNAL Patent: US 6734007-A 5 11-MAY-2004;
FEATURES
source
1. .4042
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 3675; DB 6; Length 4042;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3675; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGCCCGCGCGCGGATGGCGGGCTGGCGCGCTGTGGATGCTGCTGGCGCAG 60
|||||
```

```
Db 99 ATGAGCCCGCGCGCGGATGGCGGGGCTGGCGCGCGCTGTGGATGCTGCTGGCGCAG 158
Qy 61 GTGGCCGAGCAGGACACTGCTGTCGCGCCATGCGGACCCGCGCAGCGCCTGGGAGCCCG 120
Db 159 GTGGCCGAGCAGGACACTGCTGTCGCGCCATGCGGACCCGCGCAGCGCCTGGGAGCCCG 218
Qy 121 AGGCTCCCGGTCCTCTCCACCCCGGAGCGCGCCGCGGCTGGATGGAAGGGCGGAATAT 180
Db 219 AGGCTCCCGGTCCTCTCCACCCCGGAGCGCGCCGCGGCTGGATGGAAGGGCGGAATAT 278
Qy 181 GACCTGCTCTGCTGCTTACGAGGTTGACACAGGGGCGATTTACGTGTCCCATGAAATCATG 240
Db 279 GACCTGCTCTGCTGCTTACGAGGTTGACACAGGGGCGATTTACGTGTCCCATGAAATCATG 338
Qy 241 CACCATCAGCGCGGAGAAAGAGCAGTGGCGCTGTCCGAGGTTGAGTCTCTTTCACCTTCGG 300
Db 339 CACCATCAGCGCGGAGAAAGAGCAGTGGCGCTGTCCGAGGTTGAGTCTCTTTCACCTTCGG 398
Qy 301 CTGAAAGGCTCCAGGCAACGACTTCCAGCTGGATCTGAGGACTTCCAGGAGCCTAGTGGCT 360
Db 399 CTGAAAGGCTCCAGGCAACGACTTCCAGCTGGATCTGAGGACTTCCAGGAGCCTAGTGGCT 458
Qy 361 CTGCGCTTTATTGTGCAGACGTTGGGAAAGACAGGCACTAAGTCTGTGCAGACTTTTACCG 420
Db 459 CTGCGCTTTATTGTGCAGACGTTGGGAAAGACAGGCACTAAGTCTGTGCAGACTTTTACCG 518
Qy 421 CCAGAGGACTTCTGTTTCTTATCAAGGCTCTTTGCCGATCACACAGAAATCTCCTCAGTGGCC 480
Db 519 CCAGAGGACTTCTGTTTCTTATCAAGGCTCTTTGCCGATCACACAGAAATCTCCTCAGTGGCC 578
Qy 481 CTTTCAACCTGCAAGGCTTGTGAGGATGATAAGAAAGAGAGGAGGAGGAGTACTTCTCTA 540
Db 579 CTTTCAACCTGCAAGGCTTGTGAGGATGATAAGAAAGAGAGGAGGAGGAGTACTTCTCTA 638
Qy 541 AGGCCACTTCTTCAACCTCTCATGGAATCTCGGCAGAGCTGCCCAAGGAGCAGCTCGCCA 600
Db 639 AGGCCACTTCTTCAACCTCTCATGGAATCTCGGCAGAGCTGCCCAAGGAGCAGCTCGCCA 698
Qy 601 TCCGAGTACTGTACAAGAGATCCACAGAGCCCCATGCTCTGGGGCCAGTGAAGTCTCTG 660
Db 699 TCCGAGTACTGTACAAGAGATCCACAGAGCCCCATGCTCTGGGGCCAGTGAAGTCTCTG 758
Qy 661 GTGACTCAAGGACATGGGAGCTGGCAATCAACCCCTGCAACAGCAGGAGCCTTTCGGCTG 720
Db 759 GTGACTCAAGGACATGGGAGCTGGCAATCAACCCCTGCAACAGCAGGAGCCTTTCGGCTG 818
Qy 721 GGACTGCCACAAAGCAGCATTTCTGTGGAAGCGCAAGAAATACATGCCCCAGCCTCCC 780
Db 819 GGACTGCCACAAAGCAGCATTTCTGTGGAAGCGCAAGAAATACATGCCCCAGCCTCCC 878
Qy 781 AAGGAAAGACTCTTCAATCTTGCAGATGATATAGTCTTGTCTTACGGCATAGCGCTCT 840
Db 879 AAGGAAAGACTCTTCAATCTTGCAGATGATATAGTCTTGTCTTACGGCATAGCGCTCT 938
Qy 841 CTTCTGAGTCCCATAGAAATGAAGAACTGAACGTGGAGACCTTGTGTGTGTCGACAAA 900
Db 939 CTTCTGAGTCCCATAGAAATGAAGAACTGAACGTGGAGACCTTGTGTGTGTCGACAAA 998
Qy 901 AAGATGATGCAAAACCATGGCCATGAAATATATCCACCTAGCTGTCTCAGATACTCAAC 960
Db 999 AAGATGATGCAAAACCATGGCCATGAAATATATCCACCTAGCTGTCTCAGATACTCAAC 1058
Qy 961 ATGGTATCTGCTTTTATTCAAGATGGAACAATAGGAGGAAACATCAACATTGCAATTGTA 1020
Db 1059 ATGGTATCTGCTTTTATTCAAGATGGAACAATAGGAGGAAACATCAACATTGCAATTGTA 1118
Qy 1021 GGTCTGATTTCTTAGAAGATGAAACGCGAGGACTGGTGTGATAGTCAACGCGAGACAC 1080
Db 1119 GGTCTGATTTCTTAGAAGATGAAACGCGAGGACTGGTGTGATAGTCAACGCGAGACAC 1178
Qy 1081 ACCTTAAGTAGCTTCTGCCAGTGGCAGTGTGATGATGGGAAAGATGGGACTCTCAT 1140
Db 1179 ACCTTAAGTAGCTTCTGCCAGTGGCAGTGTGATGATGGGAAAGATGGGACTCTCAT 1238
```


Db	3399	CCAGCCTGGAGCTGGAACTGGCTTGGCCCGCTTTCATGCCCCAGGCACCCCCCATTT	3458	
Qy	3361	GCTGCTCGGGACCTCGAGGGGAGCTGGTTTGCTTCCCTCAGCTGAGCGGCC	3420	
Db	3459	GCTGCTCGGGGACCTCGAGGGGAGCTGGTTTGCTTCCCTCAGCTGAGCGGCC	3518	
Qy	3421	AGCTGCGGGGAGGGCTTGACAGAGAGTCCGTGAGTGCCTGGCTGGGGCCGCGCGGCC	3480	
Db	3519	AGCTGCGGGGAGGGCTTGACAGAGAGTCCGTGAGTGCCTGGCTGGGGCCGCGCGGCC	3578	
Qy	3481	TCAGGCTGCCTCTGACACAGAGCCCTTGGCCCTCCCTGCGCTGCACACTCACTTCTGC	3540	
Db	3579	TCAGGCTGCCTCTGACACAGAGCCCTTGGCCCTCCCTGCGCTGCACACTCACTTCTGC	3638	
Qy	3541	CCCAATTCAGAGAGAGAGTGGCTTCTGCAAGAGTACTTCCACTGGTGTACTCTGGTA	3600	
Db	3639	CCCAATTCAGAGAGAGAGTGGCTTCTGCAAGAGTACTTCCACTGGTGTACTCTGGTA	3698	
Qy	3601	CCCAGACCGGATGTGACGCCCAAGTCTTACGGCAGCAGTGTGCAAGACTTGTCTCT	3660	
Db	3699	CCCAGACCGGATGTGACGCCCAAGTCTTACGGCAGCAGTGTGCAAGACTTGTCTCT	3758	
Qy	3661	AGCTCCAACTTGA 3675		
Db	3759	AGTCCAACTTGTGA 3773		
RESULT 5				
LOCUS	BD170564	3675 bp	DNA	linear
DEFINITION	Novel protease.			PAT 17-JAN-2003
ACCESSION	BD170564			
VERSION	BD170564.1	GI:2876376		
KEYWORDS	WO 02051998-A/1.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Yamamoto; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1 (bases 1 to 3675)			
TITLE	Yamaji, N., Nishimura, K., Abe, K. and Ogino, M.			
JOURNAL	Novel protease			
COMMENT	Patent: WO 02051998-A 1 04-JUL-2002; YAMANOCHI PHARMACEUTICAL CO LTD, NOBORU YAMAJI, KOICHI NISHIMURA, KUNITAKE ABE, MAKOTO OGINO OS Homo sapiens (human) PN WO 02051998-A/1 PD 04-JUL-2002 PF 21-DEC-2001 WO 2001JP011251 PR 25-DEC-2000 JP 00P 393372 PI NOBORU YAMAJI, KOICHI NISHIMURA, KUNITAKE ABE, MAKOTO OGINO PC C12N15/09, C07K16/40, C12N5/10, C12N9/50, G01N33/15, G01N33/50, PC G01N33/573// PC C12P21/08, C12Q1/37 CC Novel protease FT Key Location/Qualifiers			
FEATURES	source	1..3675	Location/Qualifiers	
ORIGIN				
Query Match	100.0%;	Score 3673.4;	DB 6;	Length 3675;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3674;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	ATGAAGCCCGCGCGGATGGCGGGCTTGGCGCGCTGGATGCTGCTGGCGCGAG	60	
Db	1	ATGAAGCCCGCGCGGATGGCGGGCTTGGCGCGCTGGATGCTGCTGGCGCGAG	60	
Qy	61	GTGGCCGAGCAGGCACCTGCTGCGCCATGGGAGCCCGCAGCGCGCTGGGAGCCCG	120	

Db	61	GTGGCCGAGCAGGCACCTGCTGCGCCATGGGAGCCCGCAGCGCGCTGGGAGCCCG	120	
Qy	121	AGCGTCCCGCGTCTCTCCACCCGCGGAGCGCCCGCGCTGGATGGAAGAGGCGAATAT	180	
Db	121	AGCGTCCCGCGTCTCTCCACCCGCGGAGCGCCCGCGCTGGATGGAAGAGGCGAATAT	180	
Qy	181	GACCTGGTCTCTGCTTACGAGGTTGACCAAGGGGGGATTTACCGTGTCCCAATAATG	240	
Db	181	GACCTGGTCTCTGCTTACGAGGTTGACCAAGGGGGGATTTACCGTGTCCCAATAATG	240	
Qy	241	CACCATCAGGGGAGAGAGCAGTGGCGCTGTCGAGGTTGAGTCTCTTTCACCTTCGG	300	
Db	241	CACCATCAGGGGAGAGAGCAGTGGCGCTGTCGAGGTTGAGTCTCTTTCACCTTCGG	300	
Qy	301	CTGAAAGGCTCCAGGCAACGACTTCCAGTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT	360	
Db	301	CTGAAAGGCTCCAGGCAACGACTTCCAGTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT	360	
Qy	361	CTGGGCTTTATGTGAG	420	
Db	361	CTGGGCTTTATGTGAG	420	
Qy	421	CCAGAGAGCTTCTGTTTCTATCAAGGCTCTTTGGGATCAACAGAGAGAGAGAGAGAGAG	480	
Db	421	CCAGAGAGCTTCTGTTTCTATCAAGGCTCTTTGGGATCAACAGAGAGAGAGAGAGAGAG	480	
Qy	481	CTTTCAACCTGCCAAGGCTTGTGAGGATGATACGAAAGAGAGAGAGAGAGAGAGAGAGAG	540	
Db	481	CTTTCAACCTGCCAAGGCTTGTGAGGATGATACGAAAGAGAGAGAGAGAGAGAGAGAGAG	540	
Qy	541	AGGCCACTTCTCTTCAACCTCTCATGGAACTCGGAGAGAGAGAGAGAGAGAGAGAGAGAG	600	
Db	541	AGGCCACTTCTCTTCAACCTCTCATGGAACTCGGAGAGAGAGAGAGAGAGAGAGAGAGAG	600	
Qy	601	TCCACGATCTGTACAGAGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	660	
Db	601	TCCACGATCTGTACAGAGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	660	
Qy	661	GTGACCTCAAGGACATGGGAGCTGGGAGCTCAACCTCTGAGAGAGAGAGAGAGAGAGAGAG	720	
Db	661	GTGACCTCAAGGACATGGGAGCTGGGAGCTCAACCTCTGAGAGAGAGAGAGAGAGAGAGAG	720	
Qy	721	GGACTGCCACAAAGCAGCATTTCTGTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	780	
Db	721	GGACTGCCACAAAGCAGCATTTCTGTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	780	
Qy	781	AAGGAAGACCTCTTTCATCTTGCAGATGAGTATAGTCTTGTTCAGGATAGAGAGAGAGAGAG	840	
Db	781	AAGGAAGACCTCTTTCATCTTGCAGATGAGTATAGTCTTGTTCAGGATAGAGAGAGAGAGAG	840	
Qy	841	CTTCTGAGGTCCCATAGAAATGAAGAACTGAACGTGGAGAGAGAGAGAGAGAGAGAGAGAGAG	900	
Db	841	CTTCTGAGGTCCCATAGAAATGAAGAACTGAACGTGGAGAGAGAGAGAGAGAGAGAGAGAGAG	900	
Qy	901	AAGATGATGCAAAACCATGCCATGAAATATACCACTTACGATGCTGCTGCTGCTGCTGCTGCTG	960	
Db	901	AAGATGATGCAAAACCATGCCATGAAATATACCACTTACGATGCTGCTGCTGCTGCTGCTGCTG	960	
Qy	961	ATGGTATCTGCTTTTATTTCAAGATGGAACATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1020	
Db	961	ATGGTATCTGCTTTTATTTCAAGATGGAACATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1020	
Qy	1021	GGTCTGATTTCTTCTAGAGATGAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1080	
Db	1021	GGTCTGATTTCTTCTAGAGATGAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1080	
Qy	1081	ACCTTAAGTAGCTTCTGCGCAGTGGAGTCTGGAATGATGGGAGAGAGAGAGAGAGAGAGAGAGAG	1140	
Db	1081	ACCTTAAGTAGCTTCTGCGCAGTGGAGTCTGGAATGATGGGAGAGAGAGAGAGAGAGAGAGAGAG	1140	
Qy	1141	GAACGAGCCTTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1200	

QY 3361 GCTGCTGCGGGACCCCTCGAGGGGACGCTGGTTTGCCCTCACCCCTGCTCAGTGCACGGCC 3420
Db 3361 GCTGCTGCGGGACCCCTCGAGGGGACGCTGGTTTGCCCTCACCCCTGCTCAGTGCACGGCC 3420
QY 3421 AGCTGTGGGGAGGCGCTTCAGACGAGGTCGTCAGTGCCTGGGCGCGCGCGCGCC 3480
Db 3421 AGCTGTGGGGAGGCGCTTCAGACGAGGTCGTCAGTGCCTGGGCGCGCGCGCGCC 3480
QY 3481 TCAGGCTGCTCTCTGCACAGAAAGCCTTCGGCCCTCCCTGGCCTGGAACACTCACTTCTGC 3540
Db 3481 TCAGGCTGCTCTCTGCACAGAAAGCCTTCGGCCCTCCCTGGCCTGGAACACTCACTTCTGC 3540
QY 3541 CCCATTGCAGAGAAAGATGCTCTTCGAAGAGCTACTTCCACATGCTGCTACCTTGTA 3600
Db 3541 CCCATTGCAGAGAAAGATGCTCTTCGAAGAGCTACTTCCACATGCTGCTACCTTGTA 3600
QY 3601 CCCAGCAGGGATGTGCAGCCACAAAGTCTACGGCAAGCAGTGTGCAAGACTTCTCT 3660
Db 3601 CCCAGCAGGGATGTGCAGCCACAAAGTCTACGGCAAGCAGTGTGCAAGACTTCTCT 3660
QY 3661 AAGTCCAACTTGTGA 3675
Db 3661 AAGTCCAACTTGTGA 3675

RESULT 6
AX319852
LOCUS AX319852 3675 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 16 from Patent WO0183782.
ACCESSION AX319852
VERSION AX319852.1 GI:17901442
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Plowman,G.D., Whyte,D., Sudarsanam,S., Manning,G., Caenepeel,S. and Payne,V.
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Novel proteases
JOURNAL Patent: WO 0183782-A 16 08-NOV-2001;
Sugen, Inc. (US)
FEATURES
source
1..3675
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.9%; Score 3671.8; DB 6; Length 3675;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3673; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAAGCCCGCGCGCGCGATGGCGGCGCTGGCGGCGCTGTGGATGCTGTGGCGCAG 60
Db 1 ATGAAGCCCGCGCGCGCGATGGCGGCGCTGGCGGCGCTGTGGATGCTGTGGCGCAG 60
QY 61 GTGGCGGAGCAGCAGCTCGTGTGGCGCATGGGACCGCGAGCGCGCGCTGGGAGCCG 120
Db 61 GTGGCGGAGCAGCAGCTCGTGTGGCGCATGGGACCGCGAGCGCGCGCTGGGAGCCG 120
QY 121 AGCGTCCCGGCTCTCTCCACCGCGGAGCGCGCGCTGGATGGAAAGGCGCAATAT 180
Db 121 AGCGTCCCGGCTCTCTCTCCACCGCGGAGCGCGCGCTGGATGGAAAGGCGCAATAT 180
QY 181 GACCTGGTCTGCTTACGAGGTGACACAGGGCGATTAAGTGCCCATGAAATCATG 240
Db 181 GACCTGGTCTGCTTACGAGGTGACACAGGGCGATTAAGTGCCCATGAAATCATG 240
QY 241 CACCATCAGCGGCGGAGAGAGCAGTGGCGGCTGTCGAGGTTGAGTCTCTTACCTTCGG 300
Db 241 CACCATCAGCGGCGGAGAGAGCAGTGGCGGCTGTCGAGGTTGAGTCTCTTACCTTCGG 300

QY 301 CTGTAAGGCTCCAGGACGACCTTCCAGCTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360
Db 301 CTGTAAGGCTCCAGGACGACCTTCCAGCTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360
QY 361 CTTGGCTTTATTTGTGACAGCCTTGGGAAAGACAGGCACTTAAGTCTCTGACAGACTTTACCG 420
Db 361 CTTGGCTTTATTTGTGACAGCCTTGGGAAAGACAGGCACTTAAGTCTCTGACAGACTTTACCG 420
QY 421 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGCGATCAACAGAAAATCTCTCAGTGGCC 480
Db 421 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGCGATCAACAGAAAATCTCTCAGTGGCC 480
QY 481 CTTTCAACCTGCGAAGGCTTGTGAGCATGATCAAGAACAGAGAGAGCAGATTACTTCTTA 540
Db 481 CTTTCAACCTGCGAAGGCTTGTGAGCATGATCAAGAACAGAGAGAGCAGATTACTTCTTA 540
QY 541 AGGCCACTTCTTTCACACCTCTCATGGAAAATCGGACAGAGCTGCCAAAGCAGCTCGCCA 600
Db 541 AGGCCACTTCTTTCACACCTCTCATGGAAAATCGGACAGAGCTGCCAAAGCAGCTCGCCA 600
QY 601 TCCCAAGTACTGTACAAGAGATCCACAGAGCCCCATGCTCTGGGGCCAGTGAAGTCTG 660
Db 601 TCCCAAGTACTGTACAAGAGATCCACAGAGCCCCATGCTCTGGGGCCAGTGAAGTCTG 660
QY 661 GTGACCTCAAGGACATGGGAGCTGGGACATCAACCCCTGCAAGCAGAGCCTTGGCCTG 720
Db 661 GTGACCTCAAGGACATGGGAGCTGGGACATCAACCCCTGCAAGCAGAGCCTTGGCCTG 720
QY 721 GGACTGCCACAAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCAAGCTCCC 780
Db 721 GGACTGCCACAAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCAAGCTCCC 780
QY 781 AAGGAAGACTCTTCAATCTTGCAGATGATATAAGTCTTGTCTAAGCATTAAGCCTCT 840
Db 781 AAGGAAGACTCTTCAATCTTGCAGATGATATAAGTCTTGTCTAAGCATTAAGCCTCT 840
QY 841 CTTCTGAGTCCATAGAAATGAAGAACTGAACCTGAGAGACCTTGGTGGTGGCAAA 900
Db 841 CTTCTGAGTCCATAGAAATGAAGAACTGAACCTGAGAGACCTTGGTGGTGGCAAA 900
QY 901 AAGATGATGCAAAACCATGGCCATGAAAATATCACCCACTAGCTGTCTCAGCATCTCAAC 960
Db 901 AAGATGATGCAAAACCATGGCCATGAAAATATCACCCACTAGCTGTCTCAGCATCTCAAC 960
QY 961 ATGGTATCTGCTTTATTCAAGATGGAACAATAGGAGGAAACATCAACATTTGCAATTGTA 1020
Db 961 ATGGTATCTGCTTTATTCAAGATGGAACAATAGGAGGAAACATCAACATTTGCAATTGTA 1020
QY 1021 GGTCTGATTTCTTAGAAGATGAACAGCAGGACTGGTGTATAGTCAACAGCAGACAC 1080
Db 1021 GGTCTGATTTCTTAGAAGATGAACAGCAGGACTGGTGTATAGTCAACAGCAGACAC 1080
QY 1081 ACCTTAAGTAGCTTCTGCCAGTGGCAGCTGTGATTCATGGGGAAGATGGGACTCTCAT 1140
Db 1081 ACCTTAAGTAGCTTCTGCCAGTGGCAGCTGTGATTCATGGGGAAGATGGGACTCTCAT 1140
QY 1141 GACCAAGCCTCTTACTGACTGTCTGGAATATGTTCTCTGGAAGATGAGCCCTGTGAC 1200
Db 1141 GACCAAGCCTCTTACTGACTGTCTGGAATATGTTCTCTGGAAGATGAGCCCTGTGAC 1200
QY 1201 ACTTTGGGATTTGGACCCCATAAAGTGTGTATAATAATATCGCAGCTGCAGGATTAAT 1260
Db 1201 ACTTTGGGATTTGGACCCCATAAAGTGTGTATAATAATATCGCAGCTGCAGGATTAAT 1260
QY 1261 GAAGATACAGGCTTGGAGCTGGCCCTTCACTTGGCCATGAGTCTGGAACACACTTTGGC 1320
Db 1261 GAAGATACAGGCTTGGAGCTGGCCCTTCACTTGGCCATGAGTCTGGAACACACTTTGGC 1320
QY 1321 ATGATTTCATGATGGAGAGGGAAACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCCT 1380
Db 1321 ATGATTTCATGATGGAGAGGGAAACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCCT 1380
QY 1381 ACATTGGGACGAGGCAATGGAGTCTTCTCTGGTCACTCCCTGACGCCGCGCATCTACAC 1440

[illegible]

Db	2461	TCCTATTAATGAGCCGAGAACTTAATCGCTACTGGACCAACCAACGAGACACTGATTGTG	2520
Qy	2521	GAGCTGCTGTTTCAGGGAAGGAACCCGGGTGTTCCTGGGAATACTCATGCTCTCGCTTG	2580
Db	2521	GAGCTGCTGTTTCAGGGAAGGAACCCGGGTGTTCCTGGGAATACTCATGCTCTCGCTTG	2580
Qy	2581	GGACCTGAGAGAGCGCCCTGCTCCAGCCCACTACACTTGGGCGCATCGTGGCTCTGAG	2640
Db	2581	GGACCTGAGAGAGCGCCCTGCTCCAGCCCACTACACTTGGGCGCATCGTGGCTCTGAG	2640
Qy	2641	TGCTCCGTGTCTCTGGGAGGGGGAAGATGACCTGTGAGAGGGCTGCTACAGAGACCTG	2700
Db	2641	TGCTCCGTGTCTCTGGGAGGGGGAAGATGACCTGTGAGAGGGCTGCTACAGAGACCTG	2700
Qy	2701	AAGTTTCAAGTAAATATATGTCTTCTGCAATCCCAAGACACGACTGTGTCAGGGGCTGGTG	2760
Db	2701	AAGTTTCAAGTAAATATATGTCTTCTGCAATCCCAAGACACGACTGTGTCAGGGGCTGGTG	2760
Qy	2761	CCTTGCAAAAGTATCTGCTGCTCCAGCTGTGTCCTGGGAACTGTGAGTGCCTCGAGT	2820
Db	2761	CCTTGCAAAAGTATCTGCTGCTCCAGCTGTGTCCTGGGAACTGTGAGTGCCTCGAGT	2820
Qy	2821	CGAAGTGTGGCGGGGTGCCCAGAGCGCCCGCTGCAGTGCAACGCGGGGTGCACTAT	2880
Db	2821	CGAAGTGTGGCGGGGTGCCCAGAGCGCCCGCTGCAGTGCAACGCGGGGTGCACTAT	2880
Qy	2881	GACTCGGAGCCAGTCCCGGCCAGCTGTGCTCAGCTGTCTCCCTCAGCAGGAGGCCC	2940
Db	2881	GACTCGGAGCCAGTCCCGGCCAGCTGTGCTCAGCTGTCTCCCTCAGCAGGAGGCCC	2940
Qy	2941	TGCAACTCTCAGAGGTGCCACCTGCAATGAGCGCCCGGGCCCTGGGACAGGTCTCACAC	3000
Db	2941	TGCAACTCTCAGAGGTGCCACCTGCAATGAGCGCCCGGGCCCTGGGACAGGTCTCACAC	3000
Qy	3001	ACCTGTGGGAAGGGGTGAGGAAGCGGGCAGTGGCTGTGAAGAGCAACCAACCTCTCGGCC	3060
Db	3001	ACCTGTGGGAAGGGGTGAGGAAGCGGGCAGTGGCTGTGAAGAGCAACCAACCTCTCGGCC	3060
Qy	3061	AGAGCGCAGCTGCTGCCCGCCAGCTGTCTGCACCTCCGAGCCCAAGCCCAAGATGATGAA	3120
Db	3061	AGAGCGCAGCTGCTGCCCGCCAGCTGTCTGCACCTCCGAGCCCAAGCCCAAGATGATGAA	3120
Qy	3121	GCCTGTCTGCTTCHAGCCTGCCAAGAGCCCAAGAGCTGCAGTGGCTGGTGTCCGCTGG	3180
Db	3121	GCCTGTCTGCTTCHAGCCTGCCAAGAGCCCAAGAGCTGCAGTGGCTGGTGTCCGCTGG	3180
Qy	3181	TCCCAGTGTCTGTGACATGTGAAGAGGAACAACAGAAAGATCTTAAATGTGCTGAA	3240
Db	3181	TCCCAGTGTCTGTGACATGTGAAGAGGAACAACAGAAAGATCTTAAATGTGCTGAA	3240
Qy	3241	AAGTATGTTTCTGGAAAGTATCGAGAGCTGGCCTCAAGAAAGTGCTCACTTGTCCGAAG	3300
Db	3241	AAGTATGTTTCTGGAAAGTATCGAGAGCTGGCCTCAAGAAAGTGCTCACTTGTCCGAAG	3300
Qy	3301	CCAGCCTGAGCTGGAACGTTGCTCGCCCGCTGCTCCATGCTCCAGGACACCCCGCATTT	3360
Db	3301	CCAGCCTGAGCTGGAACGTTGCTCGCCCGCTGCTCCATGCTCCAGGACACCCCGCATTT	3360
Qy	3361	GCTGCTCGGGACCTCTGAGGGGCGAGCTGGTTTGCTCAACCTGGTCTCAGTGCAACGGCC	3420
Db	3361	GCTGCTCGGGACCTCTGAGGGGCGAGCTGGTTTGCTCAACCTGGTCTCAGTGCAACGGCC	3420
Qy	3421	AGCTGTGGGGAGGGCTTCAGACGAGGTCCGTGCAGTGCCTGGTGGGGGCGGCGCGGCC	3480
Db	3421	AGCTGTGGGGAGGGCTTCAGACGAGGTCCGTGCAGTGCCTGGTGGGGGCGGCGCGGCC	3480
Qy	3481	TCAGGCTGCTCTCTGCAACAGAGCCTTTCGGGCTCCCTGGCTGCAACACTCTCTCTGC	3540
Db	3481	TCAGGCTGCTCTCTGCAACAGAGCCTTTCGGGCTCCCTGGCTGCAACACTCTCTCTGC	3540
Qy	3541	CCCATTTGCAGAGAGAAAGATGCCTTCTGCAAGACTACTTTTCACTGGTGTACTCTGGTA	3600
Db	3541	CCCATTTGCAGAGAGAAAGATGCCTTCTGCAAGACTACTTTTCACTGGTGTACTCTGGTA	3600

Qy	3601	CCCAGACGGGATGTGCAGCCACAAAGTTCTACCGCAAGCAGTGTCTGCAAGACTTGTCT	3660
Db	3601	CCCAGACGGGATGTGCAGCCACAAAGTTCTACCGCAAGCAGTGTCTGCAAGACTTGTCT	3660
Qy	3661	AAGTCCAACTTGTA	3675
Db	3661	AAGTCCAACTTGTA	3675
RESULT 7			
BD170080			
LOCUS			
BD170080			
DEFINITION			
Novel ADAMTS family polypeptide and gene encoding the same.			
ACCESSION			
BD170080.1			
VERSION			
GI:27875892			
KEYWORDS			
WO 0231163-A/2.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1 (bases 1 to 5610)			
AUTHORS			
Ohara,O., Nagase,T., Nomura,N., Yano,K., Murakami,K., Yasuda,S. and			
Kanzaki,K.			
TITLE			
Novel ADAMTS family polypeptide and gene encoding the same			
JOURNAL			
Patent: WO 0231163-A 2 18-APR-2002;			
KAZUSA DNA RESEARCH INSTITUTE,MITSUBISHI PHARMA CORP,OSAMU OHARA,			
TAKAHIRO NAGASE,NOBUO NOMURA,KAZUHIRO YANO,KOJI MURAKAMI,SHINICHIRO			
YASUDA,KOJI KANZAKI			
COMMENT			
OS Homo sapiens (human)			
PN WO 0231163-A/2			
PD 18-APR-2002			
PF 11-OCT-2001 WO 2001JP008913			
PR 11-OCT-2000 JP 00P 311309,02-APR-2001 JP 01P 102905 PI			
OSAMU OHARA,TAKAHIRO NAGASE,NOBUO NOMURA,KAZUHIRO YANO,KOJI PI			
MURAKAMI,			
PI SHINICHIRO YASUDA,KOJI KANZAKI			
PC C12N15/57,C12N9/64,C12N5/00,C12N1/19,C12N1/21,C12N1/15,C07K16/			
PC 40,C12Q1/68,			
PC G01N33/50			
CC Novel ADAMTS family polypeptide and gene encoding the same FH			
Key Location/Qualifiers			
FT CDS Location/Qualifiers			
source 1..5610			
/organism="Homo sapiens"			
/mol_type="genomic DNA"			
ORIGIN /db_xref="taxon:9606"			
Query Match 99.9%; Score 3671.8; DB 6; Length 5610;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 3673; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	1	ATGAAGCCCGCGCGCGGATGCGGGGCTTGGCGGCGCTGTGGATGTCTGTGGGCGAG	60
Db	770	ATGAAGCCCGCGCGCGGATGCGGGGCTTGGCGGCGCTGTGGATGTCTGTGGGCGAG	829
Qy	61	GTGGCCGAGCAGGCACCTGCGTGCATCGGACCCGACGCGAGCGCTGTGGAGCCCG	120
Db	830	GTGGCCGAGCAGGCACCTGCGTGCATCGGACCCGACGCGCGCGCTGTGGAGCCCG	889
Qy	121	AGCGTCCCGGCTCTCTCCACCGCGGAGCGCGGCTGGATGGAAAGGCGGAATAT	180
Db	890	AGCGTCCCGGCTCTCTCTCCACCGCGGAGCGCGGCTGGATGGAAAGGCGGAATAT	949
Qy	181	GACCTGTCTCTGCTACGAGGTTGACACAGAGGGCGGATTAACGTGCCATGAATCATG	240
Db	950	GACCTGTCTCTGCTACGAGGTTGACACAGAGGGCGGATTAACGTGCCATGAATCATG	1009
Qy	241	CACCATCAGCGGCGGAGAGAGCAGTGGCCGTGTCCGAGGTTGAGTCTCTTCACTTCG	300
Db	1010	CACCATCAGCGGCGGAGAGAGCAGTGGCCGTGTCCGAGGTTGAGTCTCTTCACTTCG	1069

Qy	1381	ACATTGGCAGGACGCAATGAGTCTTCTCTCTGGTCAACCTCTGAGCGCCAGTATCTACAC	1440
Db	2150	ACATTGGCAGGACGCAATGAGTCTTCTCTCTGGTCAACCTCTGAGCGCCAGTATCTACAC	2209
Qy	1441	AAATTTCTTAAGCACCCGCTCAAGCTATCTCGCTTGTGATCAGCCAAAGCCTGTGAAGAA	1500
Db	2210	AAATTTCTTAAGCACCCGCTCAAGCTATCTCGCTTGTGATCAGCCAAAGCCTGTGAAGAA	2269
Qy	1501	TACAAGTATCCTGAGAAATTTGCCAGGAGAAATATATGATGCAAAACACAGTGCAAAGTGG	1560
Db	2270	TACAAGTATCCTGAGAAATTTGCCAGGAGAAATATATGATGCAAAACACAGTGCAAAGTGG	2329
Qy	1561	CAGTTCGGAGAGAAAGCCAAAGCTCTGCATGCTGGACTTTAAAAGGACATCTGTAAGCC	1620
Db	2330	CAGTTCGGAGAGAAAGCCAAAGCTCTGCATGCTGGACTTTAAAAGGACATCTGTAAGCC	2389
Qy	1621	CTGTGTGCCATCGTATTGGAAGGAAATGTGAGACTAAATATTATGCCAGCAGCAGAAGGC	1680
Db	2390	CTGTGTGCCATCGTATTGGAAGGAAATGTGAGACTAAATATTATGCCAGCAGCAGAAGGC	2449
Qy	1681	ACAAATTTGTGGCATGACATGTGGTGCAGGAGGAGACAGTGTGTGAATATPGTGATGAA	1740
Db	2450	ACAAATTTGTGGCATGACATGTGGTGCAGGAGGAGACAGTGTGTGAATATPGTGATGAA	2509
Qy	1741	GGCCCCAAGCCCAACCCATGGCCACATGGTCTGGACTGGTCTTCTTGGTCCCAATGCTCCAGG	1800
Db	2510	GGCCCCAAGCCCAACCCATGGCCACATGGTCTGGACTGGTCTTCTTGGTCCCAATGCTCCAGG	2569
Qy	1801	ACCTCGGAGGGGAGTATCTCATAGGAGTGCCTCTGCAACCAACCCCAAGCCATGCGAT	1860
Db	2570	ACCTCGGAGGGGAGTATCTCATAGGAGTGCCTCTGCAACCAACCCCAAGCCATGCGAT	2629
Qy	1861	GGAGGGAAGTTCTGTGAGGGCTTCCACTCGCACTCTGAAGCTCTGCAACAGTCAGAAATGT	1920
Db	2630	GGAGGGAAGTTCTGTGAGGGCTTCCACTCGCACTCTGAAGCTCTGCAACAGTCAGAAATGT	2689
Qy	1921	CCCCGGGACAGTGTGACTCTCGTGTGCTCAGTGTGCCAGACCAACAGCAGACGATTC	1980
Db	2690	CCCCGGGACAGTGTGACTCTCGTGTGCTCAGTGTGCCAGACCAACAGCAGACGATTC	2749
Qy	1981	AGAGGGCGGCACTACAAGTGAAGCCTTACACTCAAGTGAAGATCAGGACTTATGCMAA	2040
Db	2750	AGAGGGCGGCACTACAAGTGAAGCCTTACACTCAAGTGAAGATCAGGACTTATGCMAA	2809
Qy	2041	CTCTACTGTATCGCAGAAAGGATTTGATTTCTTTTCTTTTCTTTTCAAATAAAGTCAAAGAT	2100
Db	2810	CTCTACTGTATCGCAGAAAGGATTTGATTTCTTTTCTTTTCTTTTCAAATAAAGTCAAAGAT	2869
Qy	2101	GGGACTCCATGCTCGGAGATAGCCGTAATGTTGTATAGATGGGATATGTGAGAGATT	2160
Db	2870	GGGACTCCATGCTCGGAGATAGCCGTAATGTTGTATAGATGGGATATGTGAGAGATT	2929
Qy	2161	GGATGTGCAATGTCTTTGGATCTGATGCTGTTGAAGACGCTCTGTGGGTGTGTAAACGG	2220
Db	2930	GGATGTGCAATGTCTTTGGATCTGATGCTGTTGAAGACGCTCTGTGGGTGTGTAAACGG	2989
Qy	2221	AATAACTCAGCCTCAGCATTTCAAGGGGTCTCTACCAAGCACCAACCAACCAACCCAG	2280
Db	2990	AATAACTCAGCCTCAGCATTTCAAGGGGTCTCTACCAAGCACCAACCAACCAACCCAG	3049
Qy	2281	TATTATCACATGTTACCAATTCCTTCTCGAGCCGGAGTATCCGCATCTATGAATGAAC	2340
Db	3050	TATTATCACATGTTACCAATTCCTTCTCGAGCCGGAGTATCCGCATCTATGAATGAAC	3109
Qy	2341	GTCTCTACCTCTACATTTCTGTGGCAATGCCCCTGAGAGGTACTACTGTAATGGGCAC	2400
Db	3110	GTCTCTACCTCTCTACATTTCTGTGGCAATGCCCCTGAGAGGTACTACTGTAATGGGCAC	3169
Qy	2401	TGGACCGTGAATGCGCCCGGCGGTACAAATTTTTCGGGCACTACTTTTCGACTACAGACGG	2460
Db	3170	TGGACCGTGAATGCGCCCGGCGGTACAAATTTTTCGGGCACTACTTTTCGACTACAGACGG	3229
Qy	2461	TCCTATATGAGCCCGAGAACTTAAATCGCTACTGTGGACCAACCAAGAGACATGTAATGTG	2520

Db	3230	TCCTAATATGAGCCCGAGACTTAATCGCTACTGCAACCAACGAGACACTGATTGTG	3289
Qy	2521	GAGCTGCTGTTTTCAGGGGAAGGAAACCCGGGTGTCCTCGGGAACTCCATGCCTCGCTTG	2580
Db	3290	GAGCTGCTGTTTTCAGGGGAAGGAAACCCGGGTGTCCTCGGGAACTCCATGCCTCGCTTG	3349
Qy	2581	GGGACCGAGAAAGACGCCCTGCCCCAGCCAGCTACACTTTGGGCCATCTGCGCTCTGAG	2640
Db	3350	GGGACCGAGAAAGACGCCCTGCCCCAGCCAGCTACACTTTGGGCCATCTGCGCTCTGAG	3409
Qy	2641	TGCTCCGTGTCCTCGGAGGGGACATGATCCGTGAGAGAGGGCTGTCTACAGACCTTG	2700
Db	3410	TGCTCCGTGTCCTCGGAGGGGACATGATCCGTGAGAGAGGGCTGTCTACAGACCTTG	3469
Qy	2701	AAGTTTCAAGTAAATATGTCCTTCTGCAATCCCAAGACACGACTGTGCAGGGGCTGGTG	2760
Db	3470	AAGTTTCAAGTAAATATGTCCTTCTGCAATCCCAAGACACGACTGTGCAGGGGCTGGTG	3529
Qy	2761	CTTTGCAAAAGTATCTGCTGTCTCCAGCTGTGTCGTGGGAACTGGAGTGCCTCGAGT	2820
Db	3530	CTTTGCAAAAGTATCTGCTGTCTCCAGCTGTGTCGTGGGAACTGGAGTGCCTCGAGT	3589
Qy	2821	CGGACGTGTGCGGGGTGCCCAGAGCGCCCCGTGCAGTGCACACGCGGGGTGCACTAT	2880
Db	3590	CGGACGTGTGCGGGGTGCCCAGAGCGCCCCGTGCAGTGCACACGCGGGGTGCACTAT	3649
Qy	2881	GACTCGGAGCCAGTCCCGGCCAGCTGTGCCCTCAGCTGTCTCCCTCCAACGAGGAGGCC	2940
Db	3650	GACTCGGAGCCAGTCCCGGCCAGCTGTGCCCTCAGCTGTCTCCCTCCAACGAGGAGGCC	3709
Qy	2941	TGCAACTCTCAGAGCTGCCACCTGCATGGAGCGCGCGGCCCTTGGCAGAGTGTCTACAC	3000
Db	3710	TGCAACTCTCAGAGCTGCCACCTGCATGGAGCGCGCGGCCCTTGGCAGAGTGTCTACAC	3769
Qy	3001	ACCTGTGGGAAGGGGTGGAGGAAGCGGGCAGTGGCCTGTAAGACGACCAACCCCTCGGCC	3060
Db	3770	ACCTGTGGGAAGGGGTGGAGGAAGCGGGCAGTGGCCTGTAAGACGACCAACCCCTCGGCC	3829
Qy	3061	AGAGCGAGCTGTGCCGACGCTGTGCACTCCGAGCCCAAGCCGACGATGATGAA	3120
Db	3830	AGAGCGAGCTGTGCCGACGCTGTGCACTCCGAGCCCAAGCCGACGATGATGAA	3889
Qy	3121	GCCTGTCTGTTTCAGCGCTGCCACAAGCCCAAGAGCTGCAGTGGTGTGTCGCTGG	3180
Db	3890	GCCTGTCTGTTTCAGCGCTGCCACAAGCCCAAGAGCTGCAGTGGTGTGTCGCTGG	3949
Qy	3181	TCCAGTGTCTGTGATGATGAAAGAGGAAACAGAAAGATTTTAAATGTGCTGAA	3240
Db	3950	TCCAGTGTCTGTGATGATGAAAGAGGAAACAGAAAGATTTTAAATGTGCTGAA	4009
Qy	3241	AAGTATGTTTCTGGAAGTATCGAGAGTGGCTCAAGAGTGTCTACATTTGCCGAAG	3300
Db	4010	AAGTATGTTTCTGGAAGTATCGAGAGTGGCTCAAGAGTGTCTACATTTGCCGAAG	4069
Qy	3301	CCGAGCTGAGCTGGAACGTGCTGCGCCCGCTTCCATGCCCCAGGCAACCCCACTT	3360
Db	4070	CCGAGCTGAGCTGGAACGTGCTGCGCCCGCTTCCATGCCCCAGGCAACCCCACTT	4129
Qy	3361	GCTGCTCGGGACCTTCGAGGGGACGTGGTTTGCTCACCTGGTCTCAGTGCACGGCC	3420
Db	4130	GCTGCTCGGGACCTTCGAGGGGACGTGGTTTGCTCACCTGGTCTCAGTGCACGGCC	4189
Qy	3421	AGCTGTGGGGAGGGCTTCAGACGAGTCCGTGCAAGTGCCTGGTGGGGGCGCGCGCC	3480
Db	4190	AGCTGTGGGGAGGGCTTCAGACGAGTCCGTGCAAGTGCCTGGTGGGGGCGCGCGCC	4249
Qy	3481	TCAGGCTGCTCTGCAACGAGGCTTCGGCTCCCTGGCTGGCAACACTCATCTTCTGC	3540
Db	4250	TCAGGCTGCTCTGCAACGAGGCTTCGGCTCCCTGGCTGGCAACACTCATCTTCTGC	4309
Qy	3541	CCCATTTGACAGAGAAGATGCTTCTTGCAAGACTACTTTCCATGGTGTACTCTGTA	3600

Db	4310	CCCAATTCAGAGAGAGAGATGCCCTTCTGCAAGACTACTTCCACTGGTGTACCTGGTA	4369
Qy	3601	CCCAGACACGGATGTGCAGCCACAAGTTCTACGGCAAGCAGTGTCTGCAAGACTTGTCTCT	3660
Db	4370	CCCAGACACGGATGTGCAGCCACAAGTTCTACGGCAAGCAGTGTCTGCAAGACTTGTCTCT	4429
Qy	3661	AAGTCCAACTTGTGA	3675
Db	4430	AAGTCCAACTTGTGA	4444
RESULT 8			
BD170083			
LOCUS			
BD170083			
DEFINITION			
Novel ADAMTS family polypeptide and gene encoding the same.			
ACCESSION			
BD170083			
VERSION			
BD170083.1 GI:27875895			
KEYWORDS			
WO 0231163-A/5.			
SOURCE			
synthetic construct			
ORGANISM			
other sequences; artificial sequences.			
REFERENCE			
1 (bases 1 to 8435)			
AUTHORS			
Ohara,O., Nagase,T., Nomura,N., Yano,K., Murakami,K., Yasuda,S. and Kanzaki,K.			
TITLE			
Novel ADAMTS family polypeptide and gene encoding the same			
JOURNAL			
Patent: WO 0231163-A 5 18-APR-2002;			
KAZUSA DNA RESEARCH INSTITUTE, MITSUBISHI PHARMA CORP, OSAMU OHARA,			
TAKAHIRO NAGASE, NOBUO NOMURA, KAZUHIRO YANO, KOJI MURAKAMI, SHINICHIRO			
YASUDA, KOJI KANZAKI			
OS Artificial Sequence			
PN WO 0231163-A/5			
PD 18-APR-2002			
PF 11-OCT-2001 WO 2001JP008913			
PR 11-OCT-2000 JP 00P 311309, 02-APR-2001 JP 01P 102905 PI			
OSAMU OHARA, TAKAHIRO NAGASE, NOBUO NOMURA, KAZUHIRO YANO, KOJI PI			
MURAKAMI			
PI SHINICHIRO YASUDA, KOJI KANZAKI			
PC .C12N15/57, C12N9/64, C12N5/00, C12N1/19, C12N1/21, C12N1/15, C07K16/			
PC 40, C12Q1/68,			
PC G01N33/50			
CC Description of Artificial Sequence:base sequence of the CC			
Plasmid			
CC pFastBac1-HT-PJ01256			
FH Key			
Location/Qualifiers			
FT source			
1. .8435			
/organism='Artificial Sequence'.			
FEATURES			
source			
1. .8435			
/organism="synthetic construct"			
/mol_type="genomic DNA"			
/db_xref="taxon:32630"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 3670; Conservative			
99.8%; Score 3668.8; DB 6; Length 8435;			
99.9%; Pred. No. 0;			
Mismatches 2; Indels 0; Gaps 0;			
Qy	1	ATGAAGCCCGCGCGCGGATGCGGGCTTGGCGCGCTGTGGATGCTGTGGCGCAG	60
Db	4052	ATGAAGCCCGCGCGCGGATGCGGGCTTGGCGCGCTGTGGATGCTGTGGCGCAG	4111
Qy	61	GTGGCCGAGAGGCACTGTGTCGCGCATGGGACCCGCGCAGCGCTGTGGAGCCCG	120
Db	4112	GTGGCCGAGAGGCACTGTGTCGCGCATGGGACCCGCGCAGCGCTGTGGAGCCCG	4171
Qy	121	AGCGTCCCGGTCTCTCCACCCGCGGAGCGCGCGCTGGATGGAAGGCGGATAT	180
Db	4172	AGCGTCCCGGTCTCTCTCCACCCGCGGAGCGCGCGCTGGATGGAAGGCGGATAT	4231
Qy	181	GACCTGTCTCTGCTACGAGGTTGACACAGGGCGGATTAACCTGTCCCATGAATCATG	240
Db	4232	GACCTGTCTCTGCTACGAGGTTGACACAGGGCGGATTAACCTGTCCCATGAATCATG	4291

Qy	241	CACCATCAGCGCGGAGAGAGCAGTGGCCGCTGTCGAGGTGAGTCTCTTCACTTCGG	300
Db	4292	CACCATCAGCGCGGAGAGAGCAGTGGCCGCTGTCGAGGTGAGTCTCTTCACTTCGG	4351
Qy	301	CTGAAGGCTTCAGGACACGACTTCCAGCTGATCTGAGGACTTCCAGCAGCCTAGTGGCT	360
Db	4352	CTGAAGGCTTCAGGACACGACTTCCAGCTGATCTGAGGACTTCCAGCAGCCTAGTGGCT	4411
Qy	361	CCTGGCTTTATGTGTCAGAGCTTTGGGAAAGACAGGCACTAAGTCTGTGTCAGACTTTACCG	420
Db	4412	CCTGGCTTTATGTGTCAGAGCTTTGGGAAAGACAGGCACTAAGTCTGTGTCAGACTTTACCG	4471
Qy	421	CCAGAGACTTCTGTTTCTTATCAAGGCTCTTTGCGATCAACAGAACTCTCTCTAGTGGCC	480
Db	4472	CCAGAGACTTCTGTTTCTTATCAAGGCTCTTTGCGATCAACAGAACTCTCTCTAGTGGCC	4531
Qy	481	CTTTCAACCTGCCAAGCTTGTTCAGGATGATACGACAGCAAGAGGAGAGATTAATCTCCTA	540
Db	4532	CTTTCAACCTGCCAAGCTTGTTCAGGATGATACGACAGCAAGAGGAGAGATTAATCTCCTA	4591
Qy	541	AGGCCACTTCTTTCACACCTCTCATGGAACCTCGGAGAGCTGCCCCAAGGAGCTCGCCA	600
Db	4592	AGGCCACTTCTTTCACACCTCTCATGGAACCTCGGAGAGCTGCCCCAAGGAGCTCGCCA	4651
Qy	601	TCCACGTACTGTACAGAGATCCACAGAGCCCATGCTCTCTGGGGCCAGTGGCTCTG	660
Db	4652	TCCACGTACTGTACAGAGATCCACAGAGCCCATGCTCTCTGGGGCCAGTGGCTCTG	4711
Qy	661	GTGACCTCAAGGACATGGGAGCTGGCACATCAACCCCTGACAGCAGCAGCTTCGGCTG	720
Db	4712	GTGACCTCAAGGACATGGGAGCTGGCACATCAACCCCTGACAGCAGCAGCTTCGGCTG	4771
Qy	721	GGACTGCCACAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCGAGCTCCC	780
Db	4772	GGACTGCCACAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCGAGCTCCC	4831
Qy	781	AGGAAAGACTCTTCTATCTTCCAGATGATATAAGTCTTGTTCAGGCAATAGCGCTCT	840
Db	4832	AGGAAAGACTCTTCTATCTTCCAGATGATATAAGTCTTGTTCAGGCAATAGCGCTCT	4891
Qy	841	CTTCTGAGTCCCATAGNAATGAAGAACTGAAGCTGGAGACCTTGTGTGGTGGCAAA	900
Db	4892	CTTCTGAGTCCCATAGNAATGAAGAACTGAAGCTGGAGACCTTGTGTGGTGGCAAA	4951
Qy	901	AGATGATGCAAAACCATGGCCATGAAATATCAACCACTTACCTGTCTCAGCATCTCAAC	960
Db	4952	AGATGATGCAAAACCATGGCCATGAAATATCAACCACTTACCTGTCTCAGCATCTCAAC	5011
Qy	961	ATGGTATCTGCTTTATTCAAGATGAAACAATAGGAGGAAACATCAACATTGCAATTGTA	1020
Db	5012	ATGGTATCTGCTTTATTCAAGATGAAACAATAGGAGGAAACATCAACATTGCAATTGTA	5071
Qy	1021	GGTCTGATTTCTTAGAAGATGAACCCAGGACTGGTGTATAGTACCCAGCAGACCAC	1080
Db	5072	GGTCTGATTTCTTAGAAGATGAACCCAGGACTGGTGTATAGTACCCAGCAGACCAC	5131
Qy	1081	ACCTTAAGTAGCTTCTGCGCAGTGGCAGTCTGGAATGATGGGAAAGATGGGACTCGTCA	1140
Db	5132	ACCTTAAGTAGCTTCTGCGCAGTGGCAGTCTGGAATGATGGGAAAGATGGGACTCGTCA	5191
Qy	1141	GACCGCCCATCTTACTGACTGTGGATATATGTTCTTGGAAAGATGAGCCCTGTGAC	1200
Db	5192	GACCGCCCATCTTACTGACTGTGGATATATGTTCTTGGAAAGATGAGCCCTGTGAC	5251
Qy	1201	ACTTTGGGATTTGACCCCATAGTGAATGTGTAGTAAATATCGCAGCTGCAGGATTAAT	1260
Db	5252	ACTTTGGGATTTGACCCCATAGTGAATGTGTAGTAAATATCGCAGCTGCAGGATTAAT	5311
Qy	1261	GAAATACAGGCTTCTGGAGCTGGCTTCCATTTGCCCATGAGTCTCGACACAACCTTTGGC	1320
Db	5312	GAAATACAGGCTTCTGGAGCTGGCTTCCATTTGCCCATGAGTCTCGACACAACCTTTGGC	5371
Qy	1321	ATGATTTATGATGGAGAGGGGAAACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCT	1380

Db	5372	ATGATT	CATGAT	GGAGGAA	CAATGTT	GTAAAA	AGTCCG	AGGGCA	CAATCAT	GTC	CCCT	5431
Qy	1381	ACATT	GGCAGG	ACGAAT	GGAATCT	CTCTCT	CGGTCA	CGCTG	CAGCG	CGCAG	TATCT	CAC
Db	5432	ACATT	GGCAGG	ACGAAT	GGAATCT	CTCTCT	CGGTCA	CGCTG	CAGCG	CGCAG	TATCT	CAC
Qy	1441	AAATTT	CTTAAG	CACCG	CTCAAG	CTATCT	CGCTTG	CTGAT	CAGCA	AGCC	TGTGA	GAA
Db	5492	AAATTT	CTTAAG	CACCG	CTCAAG	CTATCT	CGCTTG	CTGAT	CAGCA	AGCC	TGTGA	GAA
Qy	1501	TACAA	GATAT	CTCTGA	GAATTT	GGCAG	GAGAA	TTATAT	GATG	CAACA	CACAG	TGCAG
Db	5552	TACAA	GATAT	CTCTGA	GAATTT	GGCAG	GAGAA	TTATAT	GATG	CAACA	CACAG	TGCAG
Qy	1561	CAGTT	CGGAG	AAAGCC	AAAGCT	CTGCA	TGCTG	GACTTT	AAAA	AGGAC	TCTGT	AAAGCC
Db	5612	CAGTT	CGGAG	AAAGCC	AAAGCT	CTGCA	TGCTG	GACTTT	AAAA	AGGAC	TCTGT	AAAGCC
Qy	1621	CTGTG	GTGC	ATCTG	ATTTGA	AGGAA	TGTCAG	ACTAA	TTATG	CCAG	CACAG	CAGAG
Db	5672	CTGTG	GTGC	ATCTG	ATTTGA	AGGAA	TGTCAG	ACTAA	TTATG	CCAG	CACAG	CAGAG
Qy	1681	ACAA	TTTGTG	GCA	TGACAT	GTGTG	CCGG	AGGAC	CAGTGT	GTA	AAATAT	G
Db	5732	ACAA	TTTGTG	GCA	TGACAT	GTGTG	CCGG	AGGAC	CAGTGT	GTA	AAATAT	G
Qy	1741	GGCCCC	AAAGCC	AAAGCT	CTGCA	TGCTG	GACTTT	AAAA	AGGAC	TCTGT	AAAGCC	5851
Db	5792	GGCCCC	AAAGCC	AAAGCT	CTGCA	TGCTG	GACTTT	AAAA	AGGAC	TCTGT	AAAGCC	5851
Qy	1801	ACCTG	CGGAG	GGAG	TATCT	CAT	PAGG	ATCG	CTCTG	CA	CCAA	CCCG
Db	5852	ACCTG	CGGAG	GGAG	TATCT	CAT	PAGG	ATCG	CTCTG	CA	CCAA	CCCG
Qy	1861	GGAGG	GAATCT	CTG	AGG	GTCC	CACTG	CACTG	CACTG	CACTG	CACTG	CACTG
Db	5912	GGAGG	GAATCT	CTG	AGG	GTCC	CACTG	CACTG	CACTG	CACTG	CACTG	CACTG
Qy	1921	CCCCG	GACAG	TGTTG	ACTTCC	GTG	CTGCT	CAGTGT	CCG	AGC	CAACA	CAGCAG
Db	5972	CCCCG	GACAG	TGTTG	ACTTCC	GTG	CTGCT	CAGTGT	CCG	AGC	CAACA	CAGCAG
Qy	1981	AGAGG	CGGC	ACTCA	AGTGG	AGCC	TTAC	CTCA	AGT	AGAG	ATCAG	GAT
Db	6032	AGAGG	CGGC	ACTCA	AGTGG	AGCC	TTAC	CTCA	AGT	AGAG	ATCAG	GAT
Qy	2041	CTCTA	CTGTAT	CGC	AGAG	GAATTT	GATTT	CTTCT	TTTCT	TTTCT	TTTCT	TTTCT
Db	6092	CTCTA	CTGTAT	CGC	AGAG	GAATTT	GATTT	CTTCT	TTTCT	TTTCT	TTTCT	TTTCT
Qy	2101	GGG	ACTCC	ATGCT	CGG	AGG	TAG	CCG	TAATG	TTG	TAT	AGT
Db	6152	GGG	ACTCC	ATGCT	CGG	AGG	TAG	CCG	TAATG	TTG	TAT	AGT
Qy	2161	GGATG	TGCA	ATG	CTCTG	ATCTG	ATG	AGAG	CGTCTG	TGG	GGG	TG
Db	6212	GGATG	TGCA	ATG	CTCTG	ATCTG	ATG	AGAG	CGTCTG	TGG	GGG	TG
Qy	2221	AATA	CTCAG	CTG	CAG	ATTC	ACAG	GGGTCT	CTAC	CA	CAAG	CA
Db	6272	AATA	CTCAG	CTG	CAG	ATTC	ACAG	GGGTCT	CTAC	CA	CAAG	CA
Qy	2281	TAT	TAT	CAC	ATG	TAC	CAAT	CTCT	CTG	AG	CCG	GAT
Db	6332	TAT	TAT	CAC	ATG	TAC	CAAT	CTCT	CTG	AG	CCG	GAT
Qy	2341	GTCT	CTAC	CTCT	CTAC	ATTTCT	GTG	CG	CAATG	CCCTC	AG	AG
Db	6392	GTCT	CTAC	CTCT	CTAC	ATTTCT	GTG	CG	CAATG	CCCTC	AG	AG
Qy	2401	TG	AC	CGT	GAC	TG	CG	CG	CG	CG	CG	CG

Db	6452	TGACCGTGGACTGGCCCGCGCGTGA	CAAAATTTTCGGGCAC	TACTTTTCGGACTACAGACGG	6511
Qy	2461	TCTATAATGAGCCCGAGAACTTAAT	CGCTACTGCAACCAACGAGACACT	GTATGTG	2520
Db	6512	TCCATATATGAGCCCGAGAACTTAAT	CGCTACTGCAACCAACGAGACACT	GTATGTG	6571
Qy	2521	GAGCTGCTGTTTCAGGGAAGGAA	CCCGGGTGTGCTTGGGAATAC	TCCATGTCCTCGCTTG	6571
Db	6572	GAGCTGCTGTTTCAGGGAAGGAA	CCCGGGTGTGCTTGGGAATAC	TCCATGTCCTCGCTTG	2580
Qy	2581	GGACCGGAGACGAGCCCGCTG	CCAGCCAGCTACACTTTGGGCCAT	CGTGCCCTCTGAG	2640
Db	6632	GGACCGGAGACGAGCCCGCTG	CCAGCCAGCTACACTTTGGGCCAT	CGTGCCCTCTGAG	6691
Qy	2641	TGCTCCGTGTCTCGGAGGGGAC	GAGATGACCGGTGAGAGAGGCT	GCTACAGAGACCTG	2700
Db	6692	TGCTCCGTGTCTCGGAGGGGAC	GAGATGACCGGTGAGAGAGGCT	GCTACAGAGACCTG	6751
Qy	2701	AAGTTTCAAGTAATATGCTTCT	GTGCAATCCAAAGACGACCTGT	CAACGGGGCTGGTG	2760
Db	6752	AAGTTTCAAGTAATATGCTTCT	GTGCAATCCAAAGACGACCTGT	CAACGGGGCTGGTG	6811
Qy	2761	CTTTGCAAAGTATGCTCTGCT	CTCCAGCTGCTCGGTGGGAACT	TGGAGTGCCTGCACT	2820
Db	6812	CTTTGCAAAGTATGCTCTGCT	CTCCAGCTGCTCGGTGGGAACT	TGGAGTGCCTGCACT	6871
Qy	2821	CGGACGTGTGGCGGGGTGCC	AGAGCGCCCGTGCAGTGACAC	ACGGGGGTGCACTAT	2880
Db	6872	CGGACGTGTGGCGGGGTGCC	AGAGCGCCCGTGCAGTGACAC	ACGGGGGTGCACTAT	6931
Qy	2881	GACTCGGAGCCAGTCCCGGC	CAGCCTGTGCCCTCAGCTGCT	CCCTCAGACAGGACGCC	2940
Db	6932	GACTCGGAGCCAGTCCCGGC	CAGCCTGTGCCCTCAGCTGCT	CCCTCAGACAGGACGCC	6991
Qy	2941	TGCAACTCTCAGAGCTGCC	CACTGATGAGCGCGCGCTGG	CGCAAGAGTGTGCATCAC	3000
Db	6992	TGCAACTCTCAGAGCTGCC	CACTGATGAGCGCGCGCTGG	CGCAAGAGTGTGCATCAC	7051
Qy	3001	ACCTGTGGGAAGGGTGGAG	GAAGCGGCAAGTGGCTGT	TAAGAGACACCAACCCCTCGCC	3060
Db	7052	ACCTGTGGGAAGGGTGGAG	GAAGCGGCAAGTGGCTGT	TAAGAGACACCAACCCCTCGCC	7111
Qy	3061	AGAGCGAGCTGTGCCCG	AGCTGTGTGCACCTCCGAGCC	CAAGCCAGGATGCATGAA	3120
Db	7112	AGAGCGAGCTGTGCCCG	AGCTGTGTGCACCTCCGAGCC	CAAGCCAGGATGCATGAA	7171
Qy	3121	GCCTGTCTGTTCAGCGCT	GCACAGCCCAAGAGCTG	CGATGGCTGTCCGCTGG	3180
Db	7172	GCCTGTCTGTTCAGCGCT	GCACAGCCCAAGAGCTG	CGATGGCTGTCCGCTGG	7231
Qy	3181	TCCAGTGTCTGTGACAT	GTGAAGAGGAAACAGAAAAG	ATTTCTTAAAAATGTGCTGAA	3240
Db	7232	TCCAGTGTCTGTGACAT	GTGAAGAGGAAACAGAAAAG	ATTTCTTAAAAATGTGCTGAA	7291
Qy	3241	AAGTATGTTTCTGGAAAGT	ATCGAGCTGGCTCAAGAAAGT	GTGCATATTTGCCGAG	3300
Db	7292	AAGTATGTTTCTGGAAAGT	ATCGAGCTGGCTCAAGAAAGT	GTGCATATTTGCCGAG	7351
Qy	3301	CCCAGCTGAGTGGAAAG	CTGCTCGCGCCCGCTTCC	ATGCCCCAGGACCCCCCAATTT	3360
Db	7352	CCCAGCTGAGTGGAAAG	CTGCTCGCGCCCGCTTCC	ATGCCCCAGGACCCCCCAATTT	7411
Qy	3361	GCCTGTCCGGACCTTCG	AGGGGACGTGGTTTGCT	CACCTGGTCTCAGTGACGCGCC	3420
Db	7412	GCCTGTCCGGACCTTCG	AGGGGACGTGGTTTGCT	CACCTGGTCTCAGTGACGCGCC	7471
Qy	3421	AGCTGTGGGGAGGGCTT	CAGACGAGGTCCGTG	CAGTGCCTGGGGCCGCGCGCC	3480
Db	7472	AGCTGTGGGGAGGGCTT	CAGACGAGGTCCGTG	CAGTGCCTGGGGCCGCGCGCC	7531
Qy	3481	TCAAGCTGCTCTGTG	ACAGAGCCTTTCGGCCT	CCCTGGCCTGCAACACTCACTTCTGC	3540
Db	7532	TCAAGCTGCTCTGTG	ACAGAGCCTTTCGGCCT	CCCTGGCCTGCAACACTCACTTCTGC	7591

```

QY 3541 CCCATTGCGAGAGAAAGATGCTTCTGCAAAAGACTACTTCCACTGCTGCTACTCTGGTA 3600
DB 7592 CCCATTGCGAGAGAAAGATGCTTCTGCAAAAGACTACTTCCACTGCTGCTACTCTGGTA 7651
QY 3601 CCCACGACGGGATGTGCGACCAAGTTCTACGGCAAGCAGTGTGCAAGACTTCTGCTCT 3660
DB 7652 CCCACGACGGGATGTGCGACCAAGTTCTACGGCAAGCAGTGTGCTGCAAGACTTCTGCTCT 7711
QY 3661 AAGTCCAACCTTG 3672
DB 7712 AAGTCCAACCTTG 7723

RESULT 9
LOCUS BD170084 8505 bp DNA linear PAT 17-JAN-2003
DEFINITION Novel ADAMTS family polypeptide and gene encoding the same.
ACCESSION BD170084
VERSION BD170084.1 GI:27875896
KEYWORDS WO 0231163-A/6.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences: artificial sequences.
REFERENCE 1 (bases 1 to 8505)
AUTHORS Ohara,O., Nagase,T., Nomura,N., Yano,K., Murakami,K., Yasuda,S. and
Kanzaki,K.
TITLE Novel ADAMTS family polypeptide and gene encoding the same
JOURNAL Patent: WO 0231163-A 6 18-APR-2002;
KAZUSA DNA RESEARCH INSTITUTE,MITSUBISHI PHARMA CORP,OSAMU OHARA,
TAKAHIRO NAGASE,NOBUO NOMURA,KAZUHIRO YANO,KOJI MURAKAMI,SHINICHIRO
YASUDA,KOJI KANZAKI
OS Artificial Sequence
PN WO 0231163-A/6
PD 18-APR-2002
PF 11-OCT-2001 WO 2001JP008913
PR 11-OCT-2000 JP 00P 311309,02-APR-2001 JP 01P 102905 PI
OSAMU OHARA,TAKAHIRO NAGASE,NOBUO NOMURA,KAZUHIRO YANO,KOJI PI
MURAKAMI,
PI SHINICHIRO YASUDA,KOJI KANZAKI
PC C12N15/57,C12N9/64,C12N5/00,C12N1/19,C12N1/21,C12N1/15,C07K16/
PC 40,C12Q1/68,
PC GOIN33/50
CC Description of Artificial Sequence:base sequence of the CC
plasmid
CC pFastBac1-MS/HT-PJ01256-2
FH Key Location/Qualifiers
FT source 1..8505
FT /organism='Artificial Sequence'.
FEATURES
source
1..8505
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:12630"

ORIGIN
Query Match 99.8%; Score 3668.8; DB 6; Length 8505;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3670; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAGCCCGCGCGCGATGCGGGGCTTGGCGCGCTGTGGATGCTGTGCGCGAG 60
DB 4122 ATGAAGCCCGCGCGCGATGCGGGGCTTGGCGCGCTGTGGATGCTGTGCGCGAG 4181
QY 61 GTGGCCGAGCAGCCTGCTGCGCATGGGACCCGCGAGCGCGCTGGAGCCCG 120
DB 4182 GTGGCCGAGCAGCCTGCTGCGCATGGGACCCGCGAGCGCGCTGGAGCCCG 4241
QY 121 AGCGTCCCGGCTCTCTCCACCCGCGAGCGCGCGCTGGATGAAAGGCGCAATAT 180
DB 4242 AGCGTCCCGGCTCTCTCTCCACCCGCGAGCGCGCGCTGGATGAAAGGCGCAATAT 4301
QY 181 GACCTGGTCTCTGCTACGAGGTGTGACCAAGGGGCGATTAAGTGTCCCATGAAATCATG 240

```

```

4302 GACCTGGTCTCTGCTTACGAGGTGTGACCAAGGGGCGATTTACGTGCTCCCATGAAATCATG 4361
241 CACCATCAGCGCGGAGAGAGAGCAGTGGCGCTGTCGAGGTTGAGTCTCTTTCACCTTCGG 300
4362 CACCATCAGCGCGGAGAGAGAGCAGTGGCGCTGTCGAGGTTGAGTCTCTTTCACCTTCGG 4421
301 CTGAAAGGCTCCAGGCAAGCAGCTTCCAGTGGATCTGAGGACTTCCAGAGCCTAGTGGCT 360
4422 CTGAAAGGCGCCAGGCAAGCAGCTTCCAGTGGATCTGAGGACTTCCAGAGCCTAGTGGCT 4481
361 CTTGGGCTTTATGTCAGAGCGTTGGGAAAGACAGGCACTAAGTCTGTGAGAGCTTTACCG 420
4482 CTTGGGCTTTATGTCAGAGCGTTGGGAAAGACAGGCACTAAGTCTGTGAGAGCTTTACCG 4541
421 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTTCCGATCAACAGAAACTCTCTCAGTGGCC 480
4542 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGCGATCAACAGAAACTCTCTCAGTGGCC 4601
481 CTTTCAACCTGCAAGGCTTGTGAGGATGATACGAAAGAGAGGAGGAGGAGGAGGAGGAGG 540
4602 CTTTCAACCTGCAAGGCTTGTGAGGATGATACGAAAGAGAGGAGGAGGAGGAGGAGGAGG 4661
541 AGGCCACTTCTCTTCAACCTCTCATGGAAGCTCGGAGAGCTGCCCAAGGAGCTGCCCA 600
4662 AGGCCACTTCTCTTCAACCTCTCATGGAAGCTCGGAGAGCTGCCCAAGGAGCTGCCCA 4721
601 TCCACGCTACTGTACAAAGAGATCCACAGAGCGCCATGCTCTCTGGGCGCAGTGAGGCTCTG 660
4722 TCCACGCTACTGTACAAAGAGATCCACAGAGCGCCATGCTCTCTGGGCGCAGTGAGGCTCTG 4781
661 GTGACCTCAAGGACATGGGAGCTGGCAATCAACCCCTGCACAGCAGCGACTTTCGCGCTG 720
4782 GTGACCTCAAGGACATGGGAGCTGGCAATCAACCCCTGCACAGCAGCGACTTTCGCGCTG 4841
721 GGACTGCCACAAAGCAGCATTTCTGTGGAAGCGCAAGAAATACATGCCCGCAGCTCCC 780
4842 GGACTGCCACAAAGCAGCATTTCTGTGGAAGCGCAAGAAATACATGCCCGCAGCTCCC 4901
781 AAGGAAGACTCTTCAATCTTGCAGATGATATAAGTCTTGTTCAGGCAATAGCGCTCT 840
4902 AAGGAAGACTCTTCAATCTTGCAGATGATATAAGTCTTGTTCAGGCAATAGCGCTCT 4961
841 CTTCTGAGTCCCATAGAAATGAAGAACTGAACTGAGAGACCTTGTGTGGTGGTGGTGGTGG 900
4962 CTTCTGAGTCCCATAGAAATGAAGAACTGAACTGAGAGACCTTGTGTGGTGGTGGTGGTGG 5021
901 AAGATGATGCAAAACCATGCGCATGAAATATACACCTACGTCGTCTCAGGATCTCAAC 960
5022 AAGATGATGCAAAACCATGCGCATGAAATATACACCTACGTCGTCTCAGGATCTCAAC 5081
961 ATGGTATCTGCTTTTATTCAAAGATGAAACAAATAGGAGGAAACATCAACATTGCAATTGTA 1020
5082 ATGGTATCTGCTTTTATTCAAAGATGAAACAAATAGGAGGAAACATCAACATTGCAATTGTA 5141
1021 GGTCTGATTCTTCTAGAAGATGAACAGCAGGACTGGTGTATAGTCAACCGAGACCAC 1080
5142 GGTCTGATTCTTCTAGAAGATGAACAGCAGGACTGGTGTATAGTCAACCGAGACCAC 5201
1081 ACCTTAAGTAGCTTCTGCGCAGTGGCAGTCTGAGTTGATGGGGAAGATGGGACTCGTCAT 1140
5202 ACCTTAAGTAGCTTCTGCGCAGTGGCAGTCTGAGTTGATGGGGAAGATGGGACTCGTCAT 5261
1141 GACCAAGCCATCTTACTGACTGCTCTGGATATATGTTCTTGAAGAAATGAGGCCCTGTGAC 1200
5262 GACCAAGCCATCTTACTGACTGCTCTGGATATATGTTCTTGAAGAAATGAGGCCCTGTGAC 5321
1201 ACTTTGGGATTTGCAACCCATAGTGAATGTGTAGTAAATATCGCAGCTGCAGATTAAAT 1260
5322 ACTTTGGGATTTGCAACCCATAGTGAATGTGTAGTAAATATCGCAGCTGCAGATTAAAT 5381
1261 GAAGATACAGGCTCTGGAATCTGGCCTTCAACATTTGCCATGAGTCTGGACACAACTTTGGC 1320

```

D	b	5382	GAGATACAGGCTTTGGACTTGGCCCTTCACATTGGCCCATAGTCTGTGGACACAACATTTTGGC	5441
Q	y	1321	ATGATTTCAATGATGAGAAGGAACAATGTGTAAAAAGTCCGAGSGCAAATCATCATGTGCCCT	1380
D	b	5442	ATGATTTCAATGATGAGAAGGAACAATGTGTAAAAGTCCGAGGGCAATCATATGTGCCCT	5501
Q	y	1381	ACATTTGGCAGGACCCAATGGAGTCTTCTCTGTGCACCCCTGCAGCGCCAGTATCTACAC	1440
D	b	5502	ACATTTGGCAGGACCCAATGGAGTCTTCTCTGTGCACCCCTGCAGCGCCAGTATCTACAC	5561
Q	y	1441	AAATTTCTAAGCACCGCTCAAGCTATCTGCCTTGCTGATCAGCCAAAGCCTGTGAAGAA	1500
D	b	5562	AAATTTCTAAGCACCGCTCAAGCTATCTGCCTTGCTGATCAGCCAAAGCCTGTGAAGAA	5621
Q	y	1501	TACAAGTATCTGAGAAATTTGCCAGAGAAATTAATGATGCAAAACACACAGTGCAGTGG	1560
D	b	5622	TACAAGTATCTGAGAAATTTGCCAGAGAAATTAATGATGCAAAACACACAGTGCAGTGG	5681
Q	y	1561	CAGTTCCGAGAGAAAGCCAGCTCTGCATCTGGACTTTAAAAGGCATCTGTAAAGCC	1620
D	b	5682	CAGTTCCGAGAGAAAGCCAGCTCTGCATCTGGACTTTAAAAGGCATCTGTAAAGCC	5741
Q	y	1621	CTGTGTGCCATCGTATTGGAAAGAAATGTGAGACTAAATTTATGCCACGACGAGAAGGC	1680
D	b	5742	CTGTGTGCCATCGTATTGGAAAGAAATGTGAGACTAAATTTATGCCACGACGAGAAGGC	5801
Q	y	1681	ACAAATTTGGGCATCACATGTGTGCCGGGAGGACGATGTGTGAANAATGTGTATGAA	1740
D	b	5802	ACAAATTTGGGCATCACATGTGTGCCGGGAGGACGATGTGTGAANAATGTGTATGAA	5861
Q	y	1741	GGCCCCAAGCCACCATGGCCACTGTGTGGACTGTCTTCTGGTCCCATGCTCCAGG	1800
D	b	5862	GGCCCCAAGCCACCATGGCCACTGTGTGGACTGTCTTCTGGTCCCATGCTCCAGG	5921
Q	y	1801	ACCTGGGAGGGGAGTATCTCATAGGAGTGCCTCTGCACCAACCCAAAGCCATGCGAT	1860
D	b	5922	ACCTGGGAGGGGAGTATCTCATAGGAGTGCCTCTGCACCAACCCAAAGCCATGCGAT	5981
Q	y	1861	GGAGGGAAGTCTGTGAGGGCTCCACTCGCACTCTGAAAGCTCTGCACAGTCAGAAATGT	1920
D	b	5982	GGAGGGAAGTCTGTGAGGGCTCCACTCGCACTCTGAAAGCTCTGCACAGTCAGAAATGT	6041
Q	y	1921	CCCCGGGACAGTGTGTGACTTCGTGTCTCAGTGTCCGAGCAACACGACGAGATTC	1980
D	b	6042	CCCCGGGACAGTGTGTGACTTCGTGTCTCAGTGTCCGAGCAACACGACGAGATTC	6101
Q	y	1981	AGAGGGCGGCATACAAGTGAAGCCTTACACTCAAGTAGAAGATCAGGACTTTATGCAA	2040
D	b	6102	AGAGGGCGGCATACAAGTGAAGCCTTACACTCAAGTAGAAGATCAGGACTTTATGCAA	6161
Q	y	2041	CTCTACTGTATCGCAGAAAGGATTTGATTTCTTTTTTTTTTTTGTCAATAAAGTCAAAGAT	2100
D	b	6162	CTCTACTGTATCGCAGAAAGGATTTGATTTCTTTTTTTTTTTTGTCAATAAAGTCAAAGAT	6221
Q	y	2101	GGGACTTCCATGCTCGGAGGATAGCCGTAAATGTTTGTATAGATGGGATATGTGAGAGTT	2160
D	b	6222	GGGACTTCCATGCTCGGAGGATAGCCGTAAATGTTTGTATAGATGGGATATGTGAGAGTT	6281
Q	y	2161	GGATGTGACAAATGCTCTTGGATCTGATGCTGTTGAAAGACGCTGTGGGGTGTGTAAACGG	2220
D	b	6282	GGATGTGACAAATGCTCTTGGATCTGATGCTGTTGAAAGACGCTGTGGGGTGTGTAAACGG	6341
Q	y	2221	AATACTCAGCCTCGCATTTACAGGGGTCTCTACCAAGCAACCAACCAACCAACCAAC	2280
D	b	6342	AATACTCAGCCTCGCATTTACAGGGGTCTCTACCAAGCAACCAACCAACCAACCAAC	6401
Q	y	2281	TATTATCACATGGTCAACATTTCTTCTGGAGCCGGAGTATCGCATCTATGAAATGAC	2340
D	b	6402	TATTATCACATGGTCAACATTTCTTCTGGAGCCGGAGTATCGCATCTATGAAATGAC	6461
Q	y	2341	GTCCTACTCTCATATTTCTGTGCGCAATGCCCTCAGAAAGTACTACTGATGGGCAC	2400
D	b	6462	GTCCTACTCTCATATTTCTGTGCGCAATGCCCTCAGAAAGTACTACTGATGGGCAC	6521

Qy	2401	TGGA	CCGTGGAC	CTGCC	CGGCGGTG	CAAAATTTTCGGGCAC	TACTATTTTCGACTAC	GACACGG	2460
Db	6522	TGAC	CCGTGGAC	CTGCC	CGGCGGTG	CAAAATTTTCGGGCAC	TACTATTTTCGACTAC	GACACGG	6581
Qy	2461	TCTTA	TATATG	AGCCCG	GAACTTAATCG	CTACTTGGACCA	ACGAGACACTG	ATTGTG	2520
Db	6582	TCTTA	TATATG	AGCCCG	GAACTTAATCG	CTACTTGGACCA	ACGAGACACTG	ATTGTG	6641
Qy	2521	GAGCT	GTCTTT	TGAGG	AAGAAACCCGG	GTGTTCCTTGGGAA	TACTTCCATGC	CTCGCTTG	2580
Db	6642	GAGCT	GTCTTT	TGAGG	AAGAAACCCGG	GTGTTCCTTGGGAA	TACTTCCATGC	CTCGCTTG	6701
Qy	2581	GGAACCG	AGACGAC	CCCCCTG	CCAGCCAGCTA	CACTTTGGGCCA	TGTTGGCTCTG	AG	2640
Db	6702	GGAACCG	AGACGAC	CCCCCTG	CCAGCCAGCTA	CACTTTGGGCCA	TGTTGGCTCTG	AG	6761
Qy	2641	TGCTCCG	GTCTCT	CGGAGGG	GACAGATGAC	CCGTGAGAGAGGG	CTGTCTACAGAGAC	CTTG	2700
Db	6762	TGCTCCG	GTCTCT	CGGAGGG	GACAGATGAC	CCGTGAGAGAGGG	CTGTCTACAGAGAC	CTTG	6821
Qy	2701	AGTTT	CAAGTAA	TATATGTCT	TCTGCAATCC	CAAGACGAC	CTGTCTAC	CGGGCTTG	2760
Db	6822	AGTTT	CAAGTAA	TATATGTCT	TCTGCAATCC	CAAGACGAC	CTGTCTAC	CGGGCTTG	6881
Qy	2761	CTTTG	CAAAAGTAT	CTGCTCT	CTCCAGCTG	GTCTGGGAA	CTGGAGTGC	CTGAGT	2820
Db	6882	CTTTG	CAAAAGTAT	CTGCTCT	CTCCAGCTG	GTCTGGGAA	CTGGAGTGC	CTGAGT	6941
Qy	2821	CGGACGT	TGCGGG	GTGCC	CAGAGCCGCC	CGTGTGACAC	CGCGGGTG	CATAT	2880
Db	6942	CGGACGT	TGCGGG	GTGCC	CAGAGCCGCC	CGTGTGACAC	CGCGGGTG	CATAT	7001
Qy	2881	GACTCG	GAGC	CAGTCC	CGGCCAGCTG	CGCTCAGCCTG	CTCCCTC	CAGACG	2940
Db	7002	GACTCG	GAGC	CACTG	TC	CGGCCAGCTG	CGCTCAGCCTG	CTCCCTC	7061
Qy	2941	TGCAACT	CTCTC	CAGAGCTG	CCCCACCTG	CATGAGCGCG	CGGCCCTCGG	CAGAGTGC	3000
Db	7062	TGCAACT	CTCTC	CAGAGCTG	CCCCACCTG	CATGAGCGCG	CGGCCCTCGG	CAGAGTGC	7121
Qy	3001	ACCTGTG	GGGAAGGG	TGGAGG	AGCGGG	CAGTGGCTGTAA	GAGACCA	ACCTCTGG	3060
Db	7122	ACCTGTG	GGGAAGGG	TGGAGG	AGCGGG	CAGTGGCTGTAA	GAGACCA	ACCTCTGG	7181
Qy	3061	AGAGCG	CAGCTG	CTGCC	CCAGC	GTGTGACCTCC	GAGCCCA	GAGATGC	3120
Db	7182	AGAGCG	CAGCTG	CTGCC	CCAGC	GTGTGACCTCC	GAGCCCA	GAGATGC	7241
Qy	3121	GCCTGT	CTGCTT	CAGCGCTG	CCCAAGCC	CAAGAGCTG	CAGTGGCTG	TCGCGCTG	3180
Db	7242	GCCTGT	CTGCTT	CAGCGCTG	CCCAAGCC	CAAGAGCTG	CAGTGGCTG	TCGCGCTG	7301
Qy	3181	TCC	CAGTGTCT	GTGACATGT	GAAAGAG	GAAACACAG	AAAGATCTTAA	ATGTGCTG	3240
Db	7302	TCC	CAGTGTCT	GTGACATGT	GAAAGAG	GAAACACAG	AAAGATCTTAA	ATGTGCTG	7361
Qy	3241	AGATAT	TGTTTCT	GGAAGTAT	CGAGAG	CTGGCCTCAA	AGAGTGTCTC	ACATTTG	3300
Db	7362	AGATAT	TGTTTCT	GGAAGTAT	CGAGAG	CTGGCCTCAA	AGAGTGTCTC	ACATTTG	7421
Qy	3301	CC	CAGCTG	GAGCTG	GAAGTGTG	CGCCCGCTTCA	TGCCCCAG	GACACCCCAT	3360
Db	7422	CC	CAGCTG	GAGCTG	GAAGTGTG	CGCCCGCTTCA	TGCCCCAG	GACACCCCAT	7481
Qy	3361	GCTG	CTGCGGA	CCCTC	GAGGG	CGCAGCTG	GTGTGCTTCA	CCCTGCTC	3420
Db	7482	GCTG	CTGCGGA	CCCTC	GAGGG	CGCAGCTG	GTGTGCTTCA	CCCTGCTC	7541
Qy	3421	AGCTGT	GGGGAG	GGCTT	CAGAC	GAGTCTG	CTGAGTGC	TGGGGCGCG	3480
Db	7542	AGCTGT	GGGGAG	GGCTT	CAGAC	GAGTCTG	CTGAGTGC	TGGGGCGCG	7601

[illegible]

Qy	2041	CNCTACTGATCGCAGAGGATTTGA	TTTCTTCTTTTCTTTTGTCAAAATAAGTCAAAAGAT	2100
Db	2041	CNCTACTGATCGCAGAGGATTTGA	TTTCTTCTTTTCTTTTGTCAAAATAAGTCAAAAGAT	2100
Qy	2101	GGGACTCCATGCTCGGAGGATAGCGCT	TAAGTTTGTATAGATGGATATGTGAGAGAGTT	2160
Db	2101	GGGACTCCATGCTCGGAGGATAGCGCT	TAAGTTTGTATAGATGGGATATGTGAGAGAGTT	2160
Qy	2161	GGATGTGACATGCTCTTTGGATCTGAT	CTGTTGAAGACGTCGTGTGGGGTGTGTAAACGGG	2220
Db	2161	GGATGTGACATGCTCTTTGGATCTGAT	CTGATGCTGTTGAAGACGTCGTGTGGGGTGTGTAAACGGG	2220
Qy	2221	AATTAACCTCAGCTGCAGATTTACAG	GGGGTCTCTACACCAAGACCCACCAACACCCAG	2280
Db	2221	AATTAACCTCAGCTGCAGATTTACAG	GGGGTCTCTACACCAAGACCCACCAACACCCAG	2280
Qy	2281	TATTAATCACATGGTCAACATTTCTT	CTGGAGCCCGGAGTATCGGCATCTATGAAATGAAC	2340
Db	2281	TATTAATCACATGGTCAACATTTCTT	CTGGAGCCCGGAGTATCGGCATCTATGAAATGAAC	2340
Qy	2341	GNCTCTACCTCCTACATTTCTGTGCG	CAATGTCCTCAGAAAGGTACTACTTGAATGGGCAC	2400
Db	2341	GNCTCTACCTCCTACATTTCTGTGCG	CAATGTCCTCAGAAAGGTACTACTTGAATGGGCAC	2400
Qy	2401	TGGACCGTGGACTGGCCCGGCGGTG	CAAAATTTTCGGGCACTACTTTTCGACATACACACGG	2460
Db	2401	TGGACCGTGGACTGGCCCGGCGGTG	CAAAATTTTCGGGCACTACTTTTCGACATACACACGG	2460
Qy	2461	TCCTTAATGAGCCCGGAGAACTTAAT	TCGTACTTGGACCAACCAACGAGACACTGATGTGTG	2520
Db	2461	TCCTTAATGAGCCCGGAGAACTTAAT	TCGTACTTGGACCAACCAACGAGACACTGATGTGTG	2520
Qy	2521	GAGCTGCTGTTTCAGGGAAGGAAAC	CCGGGTGTTGCCCTTGGGAATACTTCCATGCTCCGCTTG	2580
Db	2521	GAGCTGCTGTTTCAGGGAAGGAAAC	CCGGGTGTTGCCCTTGGGAATACTTCCATGCTCCGCTTG	2580
Qy	2581	GGGACCGAGAACGACGCCCTTGCC	AGCCAGCTACACTTTGGGCCACTCGTGGGCTCTGAG	2640
Db	2581	GGGACCGAGAACGACGCCCTTGCC	AGCCAGCTACACTTTGGGCCACTCGTGGGCTCTGAG	2640
Qy	2641	TGCTCCGTGTCCTGGAGGGGACAGAT	GATGACCGTGAGAGGGCTGTCTACAGAGACCTTG	2700
Db	2641	TGCTCCGTGTCCTGGAGGGGACAGAT	GATGACCGTGAGAGGGCTGTCTACAGAGACCTTG	2700
Qy	2701	AAGTTTCAAGTAATAATGTCTCT	GCAATCCCAAGACACGACCTGTACAGGGGCTGGTG	2760
Db	2701	AAGTTTCAAGTAATAATGTCTCT	GCAATCCCAAGACACGACCTGTACAGGGGCTGGTG	2760
Qy	2761	CCTTGCAGAGTATCTGGCTGTCTCC	AGTGCTCCGTGGGAACTTGGAGTGCCTGCAGT	2820
Db	2761	CCTTGCAGAGTATCTGGCTGTCTCC	AGTGCTCCGTGGGAACTTGGAGTGCCTGCAGT	2820
Qy	2821	CGGACGTGTGGCGGGGTGCCACAG	ACGCCCGCTGCTGACAGCGGGGTGCACCTAT	2880
Db	2821	CGGACGTGTGGCGGGGTGCCACAG	ACGCCCGCTGCTGCTGACAGCGGGGTGCACCTAT	2880
Qy	2881	GACTCGGAGCAGTCCCGGCGAGCCT	GTCCTCTGACCTCTCCCTCCAGCAGGACGAGCC	2940
Db	2881	GACTCGGAGCAGTCCCGGCGAGCCT	GTCCTCTGACCTCTCCCTCCAGCAGGACGAGCC	2940
Qy	2941	TGCAACTCTCAGAGCTGCCCACTT	GCATGGAGGCGCGGCCCTTGGGACAGTGCTCACAC	3000
Db	2941	TGCAACTCTCAGAGCTGCCCACTT	GCATGGAGGCGCGGCCCTTGGGACAGTGCTCACAC	3000
Qy	3001	ACCTGTGGGAAGGGGTGGAGGAAG	CGGGCAGTGGGCTGTAAAGACACCAACCCCTCGGCC	3060
Db	3001	ACCTGTGGGAAGGGGTGGAGGAAG	CGGGCAGTGGGCTGTAAAGACACCAACCCCTCGGCC	3060
Qy	3061	AGAGCGCAGCTGCTGCCCGACGCT	GTCTGCACTCCGAGCCCAAGCCCAAGCATGATGAA	3120
Db	3061	AGAGCGCAGCTGCTGCCCGACGCT	GTCTGCACTCCGAGCCCAAGCCCAAGCATGATGAA	3120

QY 3121 GCCTGCTGCTTCAGCGCTGCCACAGCCCAAGAGCTGCAGTGGCTGGTCTCGCCTGG 3180
DB 3121 GCCTGCTGCTTCAGCGCTGCCACAGCCCAAGAGCTGCAGTGGCTGGTCTCGCCTGG 3180
QY 3181 TCCAG 3186
DB 3181 TCCAG 3186
RESULT 11
AX786884
LOCUS 4888 bp DNA linear PAT 17-JUL-2003
DEFINITION Sequence 19 from Patent WO238744.
ACCESSION AX786884
VERSION AX786884.1 GI:32954185
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Lee, E.A., Hafalia, A.J., Yue, H., Lal, P.G., Yao, M.G., Lu, Y.,
Chawla, N.K., Warren, B.A., Lu, D.A., Baughn, M.R., Delegeane, A.M.,
Burford, N., Borowsky, M.L., Lee, S., Xu, Y., Griffin, J.A.,
Kallick, D.A., Gandhi, A.R., Arvizu, C.S., Ison, C.H., Tang, Y.T.,
Azimzai, Y., Elliott, V.S., Swarnakar, A., Ramkumar, J., Nguyen, D.B.,
Tribouley, C.M., Lo, T.P., Au-Young, J., Thangavelu, K. and Kearney, L.
PROTEASES
TITLE Patent: WO 0238744-A 19 16-MAY-2002;
JOURNAL Incyte Genomics, Inc. (US)
FEATURES
source Location/Qualifiers
1 .4888
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 3533147CB1"
ORIGIN
Query Match 85.3%; Score 3135.2; DB 6; Length 4888;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3163; Conservative 0; Mismatches 8; Indels 3; Gaps 2;
QY 502 TCAGGCATATACGAAACAGAGAGGAGGAGATTAATCTTAAAGGCCATTCCTTCAACACCTC 561
DB 541 TCAGGCATATACGAAACAGAGAGGAGGAGATTAATCTTAAAGGCCATTCCTTCAACACCTC 600
QY 562 TCATGAAACTCGGACAGCTGCCAAGCGAGCTCCACCTCCACCTACTGTACAGAGA 621
DB 601 TCATGAAACTCGGACAGCTGCCAAGCGAGCTCCACCTCCACCTACTGTACAGAGA 660
QY 622 TCACAGAGCCCATCTCTCTGGGGCCAGTGGAGTCTCTGCTCAACCTCAAGGACATGGGAG 681
DB 661 TCACAGAGCCCATCTCTCTGGGGCCAGTGGAGTCTCTGCTCAACCTCAAGGACATGGGAG 720
QY 682 CTGGCAATCAACCCCTGCACAGCAGCGACCTTTCGCTGGGAGCTGCCAACAAGCAGCAT 741
DB 721 CTGGCAATCAACCCCTGCACAGCAGCGACCTTTCGCTGGGAGCTGCCAACAAGCAGCAT 780
QY 742 TTCTGTGGAGAGCGAAGAAATACATGCCAGCTCCCAAGGAGACCTCTTCATCTTG 801
DB 781 TTCTGTGGAGAGCGAAGAAATACATGCCAGCTCCCAAGGAGACCTCTTCATCTTG 840
QY 802 CCAGATGAGTATAAGTCTTGTTCACGGCATAAGCGCTCTCTCTGAGGTCCCATAGAAAT 861
DB 841 CCAGATGAGTATAAGTCTTGTTCACGGCATAAGCGCTCTCTCTGAGGTCCCATAGAAAT 900
QY 862 GAAGAACTGAACCTGGAGACCTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 921
DB 901 GAAGAACTGAACCTGGAGACCTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 960
QY 922 CATGAAATATACACCACTACGTGCTCAGGATCTCAACATGATGATGATGATGATGATGATGATG 981
DB 961 CATGAAATATACCACTACGTGCTCAGGATCTCAACATGATGATGATGATGATGATGATGATG 1020

QY 982 GATGGAAACAATAGGAGGAAACAATCAACATTGCAATTGTAGTCTGATCTTCTTAGAAGAT 1041
DB 1021 GATGGAAACAATAGGAGGAAACAATCAACATTGCAATTGTAGTCTGATCTTCTTAGAAGAT 1080
QY 1042 GAACAGCCAGGACTGGTGAATAGTCAACAGCAGACCAACCTTAAAGTACTTCTGCCAG 1101
DB 1081 GAACAGCCAGGACTGGTGAATAGTCAACAGCAGACCAACCTTAAAGTACTTCTGCCAG 1140
QY 1102 TGGCAGTCTGGATTCATGGGGAAAGATGGGACTCGTCATGACCACGCCATCTTACTGACT 1161
DB 1141 TGGCAGTCTGGATTCATGGGGAAAGATGGGACTCGTCATGACCACGCCATCTTACTGACT 1200
QY 1162 GGTCTGGATATATGTTCTCTGGAAAGATAGCCCTGTGACACTTTGGGATTTGACCCATA 1221
DB 1201 GGTCTGGATATATGTTCTCTGGAAAGATAGCCCTGTGACACTTTGGGATTTGACCCATA 1260
QY 1222 AGTGGAAATGTAGTAAATATCGAGCTCGACGATTAATGAAGATACAGTCTTGGACTG 1281
DB 1261 AGTGGAAATGTAGTAAATATCGAGCTCGACGATTAATGAAGATACAGTCTTGGACTG 1320
QY 1282 GCCTTCACCATTTGCCATGAGTCTGGACACAACTTTGGCATGATTCATGATGGAGAAGG 1341
DB 1321 GCCTTCACCATTTGCCATGAGTCTGGACACAACTTTGGCATGATTCATGATGGAGAAGG 1380
QY 1342 AACATGTGTAAAGAGTCCGAGGGCAACATCATGTCTCCCTACATTTGGCAGACGCAATGGA 1401
DB 1381 AACATGTGTAAAGAGTCCGAGGGCAACATCATGTCTCCCTACATTTGGCAGACGCAATGGA 1440
QY 1402 GTCTTCTCTGCTGCTGACCTGCGAGCGCCAGTATCTACACAAATTTCTAAGCACCCTCAA 1461
DB 1441 GTCTTCTCTGCTGCTGACCTGCGAGCGCCAGTATCTACACAAATTTCTAAGCACCCTCAA 1500
QY 1462 GCTATCTGCTTGTGATCAGCAAGCCCTGTGAAGGAATACAAATGATCTCTAGAAATTTG 1521
DB 1501 GCTATCTGCTTGTGATCAGCAAGCCCTGTGAAGGAATACAAATGATCTCTAGAAATTTG 1560
QY 1522 CCAGGAGAAATATATGATGACAAACAACATGTCAGTGCAGTTCGGAGAGAAAGCCAAAG 1581
DB 1561 CCAGGAGAAATATATGATGACAAACAACATGTCAGTGCAGTTCGGAGAGAAAGCCAAAG 1620
QY 1582 CTCCTGCTGCTGACTTTTAAAGGACATCTGTAAAGCCCTGTGGTCCATCGTATTGGA 1641
DB 1621 CTCCTGCTGCTGACTTTTAAAGGACATCTGTAAAGCCCTGTGGTCCATCGTATTGGA 1680
QY 1642 AGCAATGTGAGACTAAATTTATGCGAGCAGCAAGGACCAATTTGTGGGCATGACATG 1701
DB 1681 AGCAATGTGAGACTAAATTTATGCGAGCAGCAAGGACCAATTTGTGGGCATGACATG 1740
QY 1702 TGGTCCCGGGAGGACAGTGTGTGAATAATGTTGATGAAGGCCCCCAAGCCCAATGGC 1761
DB 1741 TGGTCCCGGGAGGACAGTGTGTGAATAATGTTGATGAAGGCCCCCAAGCCCAATGGC 1800
QY 1762 CACTGCTGGACTGCTTCTTGGTCCCATGCTCCAGACCTGCGAGGGGGAGTATCT 1821
DB 1801 CACTGCTGGACTGCTTCTTGGTCCCATGCTCCAGACCTGCGAGGGGGAGTATCT 1860
QY 1822 CATAGAGTCCGCTCTGCACCAACCCCAAGCCATCGCATGGAGGGAAGTTCTGTGAGGGC 1881
DB 1861 CATAGAGTCCGCTCTGCACCAACCCCAAGCCATCGCATGGAGGGAAGTTCTGTGAGGGC 1920
QY 1882 TCCACTCGCACTCTGAAGCTCTGCAACAGTCAAGAAATGTCCCCGGGACAGTGTGTGCTTC 1941
DB 1921 TCCACTCGCACTCTGAAGCTCTGCAACAGTCAAGAAATGTCCCCGGGACAGTGTGTGCTTC 1980
QY 1942 CGTGTCTCAGTGTGCGAGCACAACAGCAGACGATTCAGAGGGCGGCATCTACAGTGG 2001
DB 1981 CGTGTCTCAGTGTGCGAGCACAACAGCAGACGATTCAGAGGGCGGCATCTACAGTGG 2040
QY 2002 AAGCCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAACTCTACTGTATCGCAGAGGA 2061
DB 2041 AAGCCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAACTCTACTGTATCGCAGAGGA 2100

Qy	2062	T T T G A T T C T T C T T T T C T T T G T C A A A P A A A G T C A A A G T G G G A C T F C C A T G C T C G A G A T	2121
Db	2101	T T T G A T T C T T C T T T T C T T T G T C A A A T A A A G T C A A A G T G G G A C T C C A T G C T C G A G A T	2160
Qy	2122	A G C C G T A A T G T T T G T A T A G A T G G G A T A C T G A G A G A G T T G G A T G T G A C A A T G T C C T T G G A	2181
Db	2161	A G C C G T A A T G T T T G T A T A G A T G G G A T A T G T G A G A G A G T T G G A T G T G A C A A T G T C C T T G G A	2220
Qy	2182	T C T G A T G C T G T T G A A G C G T C T G T G G G T G T T A A C G G G A A T A A C T C A G C C T G C A C G A T T	2241
Db	2221	T C T G A T G C T G T G A A G C G T C T G T G G G T G T T A A C G G G A A T A A C T C A G C C T G C A C G A T T	2280
Qy	2242	C A C A G G G T C T C T A C A C C A G A C C A C C A C C A C C A C C A C C A C C A C C A C C A C C A C C A C C A T T A T A C A C A T G G T C A C C A T T	2301
Db	2281	C A C A G G G T C T C T A C A C C A A G A C C A C C A C C A C C A C C A C C A C C A C C A C C A C C A C C A C C A T T A T A C A C A T G G T C A C C A T T	2340
Qy	2302	C C T T C T G A G C C G G A G T A C C G C A C T A T G A A A T G A A C G T C T C T A C C T C C T A C A T T T C T	2361
Db	2341	C C T T C T G A G C C G G A G T A C C G C A C T A T G A A A T G A A C G T C T C T A C C T C C T A C A T T T C T	2400
Qy	2362	G T G G C A A T G C C T C A G A A G G T A C T A C T G A A T G G G C A C T G G A C G T G G A C T G G C C C G G C	2421
Db	2401	G T G G C A A T G C C T C A G A A G G T A C T A C C T G A A T G G G C A C T G G A C C G T G G A C T G G C C C G G C	2460
Qy	2422	C G G T A C A A A T T T T C G G G C A C T A C T T C G A C T A C A G A C G T C C T A T A T A G C C G A G A A C	2481
Db	2461	C G G T A C A A A T T T T C G G G C A C T A C T T C G N C T A C A G A C G G T C C T A T A T A T A G A C C C G A G A C	2520
Qy	2482	T T A A T C G C T A C T G A C C A A C C A A C G A G A C A C T G A T T G T G A G A C T G C T G T T T C A G G A A G G	2541
Db	2521	T T A A T C G C T A C T G G A C C A A C C A A C G A G A C A C T G A T T G T G A G A C T G C T G T T T C A G G A A G G	2580
Qy	2542	A A C C C G G G T T G C C T G G G A A T A C T C C A T G C C T G C T T G G G A C C G A G A C A G C C C C C T	2601
Db	2581	A A C C C G G G T T G C C T G G G A A T A C T C C A T G C C T C C A T G C C T G C T T G G G A C C G A G A C A G C C C C T	2640
Qy	2602	G C C A G C C C A G C T A C A T T T G G G C C A C T G T G C G C T C T G A G T G C T C C G T G T C C T C G G A G G	2661
Db	2641	G C C A G C C C A G C T A C A T T T G G G C C A C T G T G C G C T C T G A G T G C T C C G T G T C C T C G G A G G	2700
Qy	2662	G G A C A G A C A C C G T G A G A G A G G C T G C T A C A G A G A C C T G A A G T T C A A G T A A A T A T G T C C	2721
Db	2701	G G A C A G A C A C C G T G A G A G A G G C T G C T A C A G A G A C C T G A A G T T C A A G T A A A T A T G T C C	2760
Qy	2722	T T C T G C A A T C C C A A G A C A C A C C T G C A C G G G G T G G T G C C T T G C A A A G T A T C T G C C T G T	2781
Db	2761	T T C T G C A A T C C C A A G A C A C A C C T G T C A C G G G C T G T G C C T T G C A A A G T A T C T G C C T G T	2820
Qy	2782	C C T C C A G C T G C C G T G G G A C T G G A G T G C C T G C A G T G G A C G T G T G C G G G G T G C C	2841
Db	2821	C C T C C A G C T G T G C G T G G G A A C T G S A G T G C C T G C A G T C G S A C G T G T G S G G G G T G C C	2880
Qy	2842	C A G A C C G C C C C G T C A G T G A C A C G S C G G T G C A C T A T G A C T C G G A C C A G C T C C C G G C C	2901
Db	2881	C A G A C C G C C C C G T C A G T G A C A C G S C G G T G C A C T A T G A C T C G G A C C A G C T C C C G G A	2940
Qy	2902	A G C C T G T G C C C T C A G C T G T C C C T C A G A C A G G C A G C C T G C A A C T C T C A G A G T G C C C A	2961
Db	2941	G G C C T G T G C C C T C A G - C T G G T C C C T C C A G A C C A G C C G C C T G C A A C T C T C A G A G T G C C C A	2999
Qy	2962	C C T G A T G A G C C C G G C C C T G G G C A G T G C T C A C A C C T G T G G A A G G G T G G A G	3021
Db	3000	C C T G C A T G A G C C C G G C C C T G G G C A G A G T G C T C A C A C C T G T G G A A G G - - T G A G	3057
Qy	3022	A A G C G G C A G T G G C C T G T A A G A C A C C A A C C C C T C G G C C A G A C G A C G C A G C T G T G C C C A C	3081
Db	3058	G A A C G G G C A G T G G C C T G T A A G A C A C C A A C C C C T C G G C C A G A C G C A G C T G T G C C C A C	3117
Qy	3082	G C T G T C T G C A C C T C G A C C C A A G C C C A G A T G C A T G A G C C T G C T G C T T C A G G C C T G C	3141
Db	3118	G C T G T C T G C A C C T C G A C C C A A G C C C A G A T G C A T G A G C C T G C T G C T T C A G G C C T G C	3177
Qy	3142	C A A A G C C C A A A A G A C T G A T G C A T G G C T G G T G T C C G C C T G S T C C A G T G C T G T G A C A T G T	3201

	Db	3178	CACAAGCCCAAGAAGCTGCAGTGGTGTCGCCCTCGGTCCCAGTGGCTCTGTGACATGT	3237
	Qy	3202	GAAAGAGGAACACAGAAAAGATTCTTTAAAAATGTGCTCGAAAAGTATGTTTTCTGGAAAAGTAT	3261
	Dd	3238	GAAAGAGGAACACAGAAAAGATTCTTTAAAAATGTGCTCGAAAAGTATGTTTTCTGGAAAAGTAT	3297
	Qy	3262	CGAGAGCTGGCCTCAAAGAAAGTCTCACATTTGCGGAAGCCAGCCTGGAGCTGGACGCT	3321
	Dd	3298	CGAGAGCTGGCCTCAAAGAAAGTCTCACATTTGCGGAAGCCAGCCTGGAGCTGGAACGCT	3357
	Qy	3322	GCCTGCGCCCGCTTCCATCCCCAGGACACCCCCCATTTGCTGTCGGGACCCTTCGAGG	3381
	Dd	3358	GCCTGCGCCCGCTTCCATCCCCAGGACACCCCCCATTTGCTGTCGGGACCCTTCGAGG	3417
	Qy	3382	GGCAGCTGGTTTGGCTCACCCCTGGTCTCAGTCAGCAGCGCCAGCTGTGGGGAGAGCGTTTCAG	3441
	Dd	3418	GGCAGCTGGTTTGGCTCACCCCTGGTCTCAGTCAGTCAGCGCCAGCTGTGGGGAGAGCGTTTCAG	3477
	Qy	3442	ACGAGGTCGGTCAGTGCGCTGGCTGGGGGCGCGCGGCTTCAGGCTGCCTTCCTGCACCCAG	3501
	Dd	3478	ACGAGGTCGGTCAGTGCGCTGGCTGGGGGCGCGCGGCTTCAGGCTGCCTTCCTGCACCCAG	3537
	Qy	3502	AAGCCTTGGGCTCCCTGGGCTTGAACACTCATTCTTGCCCAATTCGAGAAGAAGAAGAT	3561
	Dd	3538	AAGCCTTGGGCTCCCTGGGCTTGAACACTCATTCTTGCCCAATTCGAGAAGAAGAAGAT	3597
	Qy	3562	GCCTTCTGCAAGACTACTTCCACTGGTGTACTGTTACCTACCGGACACCGGATGTCGAGC	3621
	Dd	3598	GCCTTCTGCAAGACTACTTCCACTGGTGTACTGTTACCGGACACCGGATGTCGAGC	3657
	Qy	3622	CACAGTTCTACGGCAAGCAGTCTCGAAGACTGCTTAAGTCCAACCTTGGA	3675
	Dd	3658	CACAGTTCTACGGCAAGCAGTCTCGAAGACTGCTTAAGTCCAACCTTGGA	3711
RESULT 12				
BD170079				
LOCUS	BD170079	Novel ADAMTS family polypeptide and gene encoding the same.	4234 bp DNA linear	PAT 17-JAN-2003
DEFINITION	BD170079			
ACCESSION	BD170079.1	GI:27875891		
VERSION	WO 0231163-A/1.			
KEYWORDS	Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Chara,O., Nagase,T., Nomura,N., Yano,K., Murakami,K., Yasuda,S. and Kanzaki,K.			
TITLE	Novel ADAMTS family polypeptide and gene encoding the same			
JOURNAL	Patent: WO 0231163-A/1 18-APR-2002;			
	KAZUSA DNA RESEARCH INSTITUTE,MITSUBISHI PHARMA CORP,OSAMU OHARA,			
	TAKAHIRO NAGASE,NOBUO NOMURA,KAZUHIRO YANO,KOJI MURAKAMI,SHINICHIRO			
	YASUDA, KOJI KANZAKI			
COMMENT	OS Homo sapiens (human)			
	PN WO 0231163-A/1			
	PD 18-APR-2002			
	PF 11-OCT-2001 WO 2001JP008913			
	PR 11-OCT-2000 JP OOP 311309,02-APR-2001 JP OIP 102905 PI			
	OSAMU OHARA, TAKAHIRO NAGASE, NOBUO NOMURA, KAZUHIRO YANO, KOJI PI			
	MURAKAMI,			
	PI SHINICHIRO YASUDA, KOJI KANZAKI			
	PC C12N15/57, C12N9/64, C12N5/00, C12N1/19, C12N1/21, C12N1/15, C07K16/			
	PC 40, C12Q1/68,			
	PC G01N33/50			
	CC Novel ADAMTS family polypeptide and gene encoding the same FH			
Key	CDS	Location/Qualifiers	(3).	(3068).
FEATURES				
source				
	1..4234			
	/organism="Homo sapiens"			
	/mol_type="genomic DNA"			

ORIGIN	/db_xref="taxon:9606"	Db
Query Match 83.4%; Score 3064.8; DB 6; Length 4234; Best Local Similarity 99.9%; Pred. No. 0; Matches 3066; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Qy 608 TACTGTACAAGAGATCCACAGAGCCCCATGCTCTCTGGGGCCAGTGAGTCTGTGTGACCT 667	1628 GCCATCGTATTGGGAAGAAATGTGAGACTAAATTTATGCCAGCAGCAGAGAAGGCCACAATTT 1687	
Db 1 TACTGTACAAGAGATCCACAGAGCCCCATGCTCTCTGGGGCCAGTGAGTCTGTGTGACCT 60	1021 GCCATCGTATTGGGAAGAAATGTGAGACTAAATTTATGCCAGCAGCAGAGAAGGCCACAATTT 1080	
Qy 668 CAGGACATGGGAGCTGGCACATCAACCCCTGCCACAGCAGCGACCTTCGGCTGGGACTGC 727	1688 GTGGGCATGACATGTGTGTGCCGGGAGGACAGTGTGTGAAATATATGTGTATGAAGGCCCA 1747	
Db 61 CAAGGACATGGGAGCTGGCACATCAACCCCTGCCACAGCAGCGACCTTCGGCTGGGACTGC 120	1081 GTGGGCATGACATGTGTGTGCCGGGAGGACAGTGTGTGAAATATATGTGTATGAAGGCCCA 1140	
Qy 728 CACAAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCCAGCGCTCCCAAGGAAG 787	1748 AGCCCCACCATGGCCACATGCTGGCATGTGCTCTTCTGTGTGCCCATGCTCCAGGACCTGCG 1807	
Db 121 CACAAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCCAGCGCTCCCAAGGAG 180	1141 AGCCCCACCATGGCCACATGCTGGCATGTGCTCTTCTGTGTGCCCATGCTCCAGGACCTGCG 1200	
Qy 788 ACCTCTTACTTTCGCAGATAGTATAAGTCTTTGCTTACCGGATAAGCGCTCTCTTTCTGA 847	1808 GAGGGGAGTATCTCATAGGAGTGGCCCTCTGCAACAAACCCCAAGCCATCGCATGGAGGGA 1867	
Db 181 ACCTCTTACTTTCGCAGATAGTATAAGTCTTTGCTTACCGGATAAGCGCTCTCTTCTGA 240	1201 GAGGGGAGTATCTCATAGGAGTGGCCCTCTGCAACAAACCCCAAGCCATCGCATGGAGGGA 1260	
Qy 848 GGTCCCATAGAAATGAGAACTGAACTGGAGCTGGAGACCTTGGTGGTGCACAAAAGATGA 907	1868 AGTTCTGTGAGGGCTCCACTGCACTCTGAAGCTCTGCAACAGTCAAGATGTCCCCCGGG 1927	
Db 241 GSTACCATGAAATGAGAACTGAACTGGAGACCTTGGTGGTGCACAAAAGATGA 300	1261 AGTTCTGTGAGGGCTCCACTGCACTCTGAAGCTCTGCAACAGTCAAGAAATGTCCCCCGGG 1320	
Qy 908 TGCAAAACATGGCCATGAAATATACACACCTAGCTGCTCACGATCTCAACATGGTAT 967	1928 ACAGTGTGACTTCCGTCGCTCAGTGTGCGGAGCAACACAGCAGACGATTCAGAGGGC 1987	
Db 301 TGCAAAACATGGCCATGAAATATACACACCTAGCTGCTCACGATCTCAACATGGTAT 360	1321 ACAGTGTGACTTCCGTCGCTCAGTGTGCGGAGCAACACGACGACGATTCAGAGGGC 1380	
Qy 968 CTGCTTTATTCAAAGATGAAACAATAGGAGGAAACATCAACATTTGCAATTGTAGTCTGA 1027	1988 GGCATCTACAAGTGGGAAGCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAACTCTTACT 2047	
Db 361 CTGCTTTATTCAAAGATGAAACAATAGGAGGAAACATCAACATTTGCAATTGTAGTCTGA 420	1381 GGCATCTACAAGTGGGAAGCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAACTCTTACT 1440	
Qy 1028 TTCTTCTAGAGATGAAACGACGACTGGTGATAGTACCAACGACACACACCTTAA 1087	2048 GTATCGCAGAGAGATTGATTTCTTTTCTTTTCTTTGTCAAAATAAGTCAAAAGATGGGACTC 2107	
Db 421 TTCTTCTAGAGATGAAACGACGAGACTGGTGATAGTACCAACGACACACACCTTAA 480	1441 GTATCGCAGAGAGATTGATTTCTTTTCTTTTCTTTGTCAAAATAAGTCAAAAGATGGGACTC 1500	
Qy 1088 GTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGGAAGATGGGACTCGTCATGACCAAG 1147	2108 CATGCTCGGAGGATAGCCGTAATGTTTGTATAGATGGGATATGTGAGAGAGTTGGATGTG 2167	
Db 481 GTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGGAAGATGGGACTCGTCATGACCAAG 540	1501 CATGCTCGGAGGATAGCCGTAATGTTTGTATAGATGGGATATGTGAGAGAGTTGGATGTG 1560	
Qy 1148 CCACTTTACTGACTGGTCTGGATATATGTTCTCGAAGAAATGAGCCCTGTGACACTTTGG 1207	2168 ACAATGTCTTGGATCTGATGCTGTGAAAGACGCTCTGTGGGGTGTGTAAACGGGAATAACT 2227	
Db 541 CCACTTTACTGACTGGTCTGGATATATGTTCTCGAAGAAATGAGCCCTGTGACACTTTGG 600	1561 ACAATGTCTTGGATCTGATGCTGTGAGAGACGCTCTGTGGGGTGTGTAAACGGGAATAACT 1620	
Qy 1208 GATTGACCCCATAGTGGAAATGTGTAGTAATATCGCAGCTGCACGANTTAATGAAGATA 1267	2228 CAGCTGCAACGATTCACAGGGGTCTCTACCAAGACACCAACCAACCAACAGATATTATC 2287	
Db 601 GATTGACCCCATAGTGGAAATGTGTAGTAATATCGCAGCTGCACGANTTAATGAAGATA 660	1621 CAGCTGCAACGATTCACAGGGGTCTCTACACCAAGACACCAACCAACAGATATTATC 1680	
Qy 1268 CAGGTCTTGGACTGGCCCTTACCAATTTGCCATGAGTCTGGACACAACTTTGGCATGATTC 1327	2288 ACATGGTCAACCATTTCTTCTTGGAGCCGGAGTATCCGATCTATGAATGAAAGTCTCTTA 2347	
Db 661 CAGGTCTTGGACTGGCCCTTACCAATTTGCCATGAGTCTGGACACAACTTTGGCATGATTC 720	1681 ACATGGTCAACCATTTCTTCTTGGAGCCGGAGTATTCGCACTATGAATGAAAGTCTCTTA 1740	
Qy 1328 ATGATGGAAGGAACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCTTACATTGG 1387	2348 CTTCTCTACATTTCTGTGGGCAATGCCCCCTCAGAAAGTACTACCTGAAATGGGCACTGGACCG 2407	
Db 721 ATGATGGAAGGAACATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCTTACATTGG 780	1741 CTTCTCTACATTTCTGTGGGCAATGCCCCCTCAGAAAGTACTACCTGAAATGGGCACTGGACCG 1800	
Qy 1388 CAGGACGCAATAGGAGTCTTCTCTGGTCAACCTCGCAGCCGCGAGTATCTACAAATTTTC 1447	2408 TGGACTGCGCCGGCCGGTACAAAATTTTGGGCACTACTTTTGGACTACAGACGCTCTTATA 2467	
Db 781 CAGGACGCAATAGGAGTCTTCTCTGGTCAACCTCGCAGCCGCGAGTATCTACAAATTTTC 840	1801 TGGACTGCGCCGGCCGGTACAAAATTTTGGGCACTACTTTTGGCACTACAGACGCTCTTATA 1860	
Qy 1448 TAAGCACCCCTCAAGCTATCTGCTTGTGATCAGCCAAAGCCCTGTGAAGAAATACAAGT 1507	2468 ATGAGCCCGAGAACTTAAATCGTACTGGAACCAACCAACGAGACACTGATTTGTGGAGCTGC 2527	
Db 841 TAAGCACCCCTCAAGCTATCTGCTTGTGATCAGCCAAAGCCCTGTGAAGAAATACAAGT 900	1861 ATGAGCCCGAGAACTTAAATCGTACTGGAACCAACCAACGAGACACTGATTTGTGGAGCTGC 1920	
Qy 1508 ATCCTGAGAAATTTGCCAGAGAAATATATGATGCAAAACACAGTGCAGTGGCAGTTTC 1567	2528 TGTTTTACGGAAGGAACCCGGGTGTGCTGGGAATACTCATGSCCTCGCTTGGGGAACCG 2587	
Db 901 ATCCTGAGAAATTTGCCAGAGAAATATATGATGCAAAACACAGTGCAGTGGCAGTTTC 960	1921 TGTTTTACGGAAGGAACCCGGGTGTGCTGGGAATACTCATGCTCTCGCTTGGGGAACCG 1980	
Qy 1568 GAGAGAAACCAAGCTCTGCATGCTGGAATTTAAAAAGGACATCTGTAAGCCCTGTGGT 1627	2588 AGAAGCAGCCCTCTGCCAGCCAGCTACATCTGGGGCCATCTGTGCGCTCTGAGTGTCTCG 2647	

Qy	788	ACCTCTTCATCTTGCAGATGAGTATTAAGTCTTGCTTACGGCATAAGCGCTCTCTTCTGA	847
Db	181	ACCTCTTCATCTTGCAGATGAGTATTAAGTCTTGCTTACGGCATAAGCGCTCTCTTCTGA	240
Qy	848	GGTCCCATAGAAATGAAAGACTGAACCTGGAGACCTTGGTGGCTCGACAAAAAGATGA	907
Db	241	GGTACCATAGAAATGAAAGACTGAACCTGGAGACCTTGGTGGCTCGACAAAAAGATGA	300
Qy	908	TGCAAAACCATGGCCATGAAATATCACCACTACGTGCTCACGATACTCAACATGGTAT	967
Db	301	TGCAAAACCATGGCCATGAAATATCACCACTACGTGCTCACGATACTCAACATGGTAT	360
Qy	968	CTGCTTTATTCAAAGATGGAAACAATAGAGGAAACATCAACATTTGGCAATTTGTAGGCTCGA	1027
Db	361	CTGCTTTATTCAAAGATGGAAACAATAGAGGAAACATCAACATTTGGCAATTTGTAGGCTCGA	420
Qy	1028	TTCTTCTAGAGATGAACAGCCAGGACTGGTGATAAGTCAACCGCAGAGCCACACCTTAA	1087
Db	421	TTCTTCTAGAGATGAACAGCCAGGACTGGTGATAAGTCAACCGCAGAGCCACACCTTAA	480
Qy	1088	GTAGCTTCTGCCAGTGGCAGTCTGGAATTGATGGGGAAGATGGGACTCGTCATGACCACG	1147
Db	481	GTAGCTTCTGCCAGTGGCAGTCTGGAATTGATGGGGAAGATGGGACTCGTCATGACCACG	540
Qy	1148	CCATCTTACTGACTGGTCTGGATATATGTTCTCTGGAAAGATGAGCCCTGTGACACTTTGG	1207
Db	541	CCATCTTACTGACTGGTCTGGATATATGTTCTCTGGAAAGATGAGCCCTGTGACACTTTGG	600
Qy	1208	GATTTGCACCCATAAGTGGATGTTAGTAAATCCAGCTGCACAGTAATTAATGACAGATA	1267
Db	601	GATTTGCACCCATAAGTGGATGTTAGTAAATCCAGCTGCACAGTAATTAATGACAGATA	660
Qy	1268	CAGTCTTTGGACTGGCTTACCATTGCCCAGTCTGGACACAACTTTGGGCATGATTC	1327
Db	661	CAGTCTTTGGACTGGCTTACCATTGCCCAGTCTGGACACAACTTTGGGCATGATTC	720
Qy	1328	ATGATGAGAAAGGAAACATGTTTAAAAGTCCGAGGGCAACATCATGTCCCTACATTTGG	1387
Db	721	ATGATGAGAAAGGAAACATGTTGCAAAAAGTCCGAGGGCAACATCATGTCCCTACATTTGG	780
Qy	1388	CAGGACCCATGGAGTCTTCTCTGGTCAACCTGCAGCGCCAGTATCTACACAAATTTTC	1447
Db	781	CAGGACCCATGGAGTCTTCTCTGGTCAACCTGCAGCGCCAGTATCTACACAAATTTTC	840
Qy	1448	TAAGCACCGCTCAAGCTATCTGCCCTGTGATCAGCCAAAGCCCTGTCAAGGAATACAAGT	1507
Db	841	TAAGCACCGCTCAAGCTATCTGCCCTGTGATCAGCCAAAGCCCTGTCAAGGAATACAAGT	900
Qy	1508	ATCCTGAGAAATTGCCAGGAAATATATGATGCAAAACACACAGTGCAGTGCAGTTTCG	1567
Db	901	ATCCTGAGAAATTGCCAGGAAATATATGATGCAAAACACACAGTGCAGTGCAGTTTCG	960
Qy	1568	GAGAGAACCCAGCTCTGCATGCTGACATTTTAAAAGGACATCTGTAAAGCCCTGTGGT	1627
Db	961	GAGAGAACCCAGCTCTGCATGCTGACATTTTAAAAGGACATCTGTGTAAAGCCCTGTGGT	1020
Qy	1628	GCCATCGTATTGGAAGGAAATGTGAGACTAAATTTATGCCAGACGAGAGGCACAATTT	1687
Db	1021	GCCATCGTATTGGAAGGAAATGTGAGACTAAATTTATGCCAGACGAGAGGCACAATTT	1080
Qy	1688	GTGGGCATGACATGTGTGCCGGGAGGACAGTGTGTGAATAATGTGTGATGAAGGCCCA	1747
Db	1081	GTGGGCATGACATGTGTGCCGGGAGGACAGTGTGTGAATAATGTGTGATGAAGGCCCA	1140
Qy	1748	AGCCCAACCCATGGCCATCTGCTCGACCAACCCCAAGCCATCGCATGGAGGGA	1807
Db	1141	AGCCCAACCCATGGCCATCTGCTCGACCAACCCCAAGCCATCGCATGGAGGGA	1200
Qy	1808	GAGGGGAGTATCTCATAGGAGTGCCTCTGACCAACCCCAAGCCATCGCATGGAGGGA	1867
Db	1201	GAGGGGAGTATCTCATAGGAGTGCCTCTGACCAACCCCAAGCCATCGCATGGAGGGA	1260
Qy	1868	AGTTCTGTGAGGGCTCCACTCGCACTCTGAAGCTCTGCAACAGTCAGAATGTGCCGGG	1927

Db	1261		AGTTCTGTGAGGGCTCCACTCGCACTCTGAAGCTCTGCAACAGTCAAGAAATGTCCTCGGG	1320
Qy	1928		ACAGTGTGACTTCCGTCGCTCAGTGTGCCGACACAAACAGACGATTCAGAGGGC	1987
Db	1321		ACAGTGTGACTTCCGTCGCTCAGTGTGCCGACAAACAGACGATTCAGAGGGC	1380
Qy	1988		GGCACTACAAAGTGGAGCCTTACACTCAAGTAGAAGTCAAGACTTATGCAAACTCTACT	2047
Db	1381		GGCACTACAAAGTGGAGCCTTACACTCAAGTAGAAGTCAAGACTTATGCAAACTCTACT	1440
Qy	2048		GTATCGCAGAGGATTTGATTTCTTTTCTTTGTGCAATAAAGTCAAAGATGGGACTC	2107
Db	1441		GTATCGCAGAGGATTTGATTTCTTTCTTTGTCAATAAAGTCAAAGATGGGACTC	1500
Qy	2108		CATGCTCGGAGGATAGCCGTAATGTTGTATAGATGCGATATCTGAGAGAGTTGGAATG	2167
Db	1501		CATGCTCGGAGGATAGCCGTAATGTTGTATAGATGCGATATCTGAGAGAGTTGGAATG	1560
Qy	2168		ACAATGTCTTGGATCTGATGCTGTTGAAGACGCTCTGTGGGGTGTGTAAACGGAAATACT	2227
Db	1561		ACAATGTCTTGGATCTGATGCTGTTGAAGACGCTCTGTGGGGTGTGTAAACGGAAATACT	1620
Qy	2228		CAGCTCGACGATTCACAGGGGTCTCTACACCAAGCAACCAACCAACCAAGTATATC	2287
Db	1621		CAGCTCGACGATTCACAGGGGTCTCTACACCAAGCAACCAACCAACCAAGTATATC	1680
Qy	2288		ACATGGTCAACATTCCTCTCGAGCCGGAGTATCGCATCTATGAAATGAACTCTCTA	2347
Db	1681		ACATGGTCAACATTCCTCTCGAGCCGGAGTATCGCATCTATGAAATGAACTCTCTA	1740
Qy	2348		CCTCTCAATTTCTGTGCGCAATGCCCTCAGAAAGTACTACCTGAATGGGCACTGACCG	2407
Db	1741		CCTCTCAATTTCTGTGCGCAATGCCCTCAGAAAGTACTACCTGAATGGGCACTGACCG	1800
Qy	2408		TGAGTGGCCCGCGGGTACAAATTTTGGGGCACTACTTTTGGCACTACAGACGTCCTATA	2467
Db	1801		TGAGTGGCCCGCGGGTACAAATTTTGGGGCACTACTTTTGGCACTACAGACGTCCTATA	1860
Qy	2468		ATGAGCCGAGAACTTAATCGCTACTGGAACCAACCAACGAGACACTGATGTTGGAGCTGC	2527
Db	1861		ATGAGCCGAGAACTTAATCGCTACTGGAACCAACCAACGAGACACTGATGTTGGAGCTGC	1920
Qy	2528		TGTTTCAGGGAAGAAACCCGGGTGTTCCCTGGGAATACTCCATGCTCGCTTTGGGACCG	2587
Db	1921		TGTTTCAGGGAAGAAACCCGGGTGTTCCCTGGGAATACTCCATGCTCGCTTTGGGACCG	1980
Qy	2588		AGAAGACGCCCCCTGCCAGCCAGCTACACTTTGGGCCATCGTGCCTCTGAGTGTCCCG	2647
Db	1981		AGAAGACGCCCCCTGCCAGCCAGCTACACTTTGGGCCATCGTGCCTCTGAGTGTCCCG	2040
Qy	2648		TGTCCTCGGAGGGGACAGATGACCGTGGAGAGGGCTGTACAGAGACCTGAGTTTC	2707
Db	2041		TGTCCTCGGAGGGGACAGATGACCGTGGAGAGGGCTGTACAGAGACCTGAGTTTC	2100
Qy	2708		AAGTAAATATGTCCTTCTGCAATCCCAAGACAGACTGTCAACGGGGCTGGTCCCTTGA	2767
Db	2101		AAGTAAATATGTCCTTCTGCAATCCCAAGACAGACTGTCAACGGGGCTGGTCCCTTGA	2160
Qy	2768		AAGTATCTGCTGTCTCCAGCTGTGCTGGGGAACTGGAGTGTGCTCGAGTGTGAGCGT	2827
Db	2161		AAGTATCTGCTGTCTCCAGCTGTGCTGGGGAACTGGAGTGTGCTCGAGTGTGAGCGT	2220
Qy	2828		GTGCGGGGGTGCCAGAGCCCGCTGACGTGACAGCGGGGTGCACTATGACTCGG	2887
Db	2221		GTGCGGGGGTGCCAGAGCCCGCTGACGTGACAGCGGGGTGCACTATGACTCGG	2280
Qy	2888		AGCCAGTCCCGGACGCTGTGCCCTCAGCCTGCTCCCTCCAGCAGGCGAGGCTGCAACT	2947
Db	2281		AGCCAGTCCCGGACGCTGTGCCCTCAGCCTGCTCCCTCCAGCAGGCGAGGCTGCAACT	2340
Qy	2948		CTCAGAGCTGCCCACTGTATGAGAGCCCGGGCCCTGGGGCAGAGTGCTCACACACTGTG	3007

Db 2341 CTGAGAGTCCACCTGCAATGAGAGCGCGCGCCCTGGGAGAGTGCTCACACACCTGTG 2400
Qy 3008 GGAAGGGGTGAGGAAGCGGGCAGTGGCTGTGAAGAGCACCAACCCCTCGGCCAGAGCGC 3067
Db 2401 GGAAGGGGTGAGGAAGCGGGCAGTGGCTGTGAAGAGCACCAACCCCTCGGCCAGAGCGC 2460
Qy 3068 AGCTGCTGCCGAGCGTGTCTGACCTCGAGGCCAAGCCAGCGATGCAATGAAGCCTGTC 3127
Db 2461 AGCTGCTGCCGAGCGTGTCTGACCTCGAGGCCAAGCCAGCGATGCAATGAAGCCTGTC 2520
Qy 3128 TGCTTCAGCGCTGCCAAGCCCAAGAGCTGAGTGGCTGGTTCGGCTGGTCCCGAGT 3187
Db 2521 TGCTTCAGCGCTGCCAAGCCCAAGAGCTGAGTGGCTGGTTCGGCTGGTCCCGAGT 2580
Qy 3188 GCTCTGTGATCATGTGAAGAGGAACACAGAAAGATTCTTAAATGTGCTGAAAAGTATG 3247
Db 2581 GCTCTGTGATCATGTGAAGAGGAACACAGAAAGATTCTTAAATGTGCTGAAAAGTATG 2640
Qy 3248 TTCTCTGAAAGTATCGAGAGCTGGCTCAAAGAGTGCTCACATTTGCCGAAGCCAGCC 3307
Db 2641 TTCTCTGAAAGTATCGAGAGCTGGCTCAAAGAGTGCTCACATTTGCCGAAGCCAGCC 2700
Qy 3308 TGGAGCTGGAACCTGCTGCGCCCGCTTCCATGCCCCAGGACACCCCATTTGCTGTG 3367
Db 2701 TGGAGCTGGAACCTGCTGCGCCCGCTTCCATGCCCCAGGACACCCCATTTGCTGTG 2760
Qy 3368 CGGAGCCCTCGAGGGGAGCTGTTTGGCTCAACCTGCTCAGTGCAAGCCAGCTGTG 3427
Db 2761 CGGAGCCCTCGAGGGGAGCTGTTTGGCTCAACCTGCTCAGTGCAAGCCAGCTGTG 2820
Qy 3428 GGGAGGCGCTTCAGAGAGCTCGTGCAGTGGCTGGCGGGCGCGGCGCTCAGGCT 3487
Db 2821 GGGAGGCGCTTCAGAGAGCTCGTGCAGTGGCTGGCGGGCGCGGCGCTCAGGCT 2880
Qy 3488 GCTCTCTGCACAGAGCCCTTGGCCCTCGCTGGCTGCAACACTCACTTCTGCCCCATTG 3547
Db 2881 GCTCTCTGCACAGAGCCCTTGGCCCTCGCTGGCTGCAACACTCACTTCTGCCCCATTG 2940
Qy 3548 CAGAGAGAAGTGTCTTGAAGACTACTTCCAGTGGTGTACTGTGTACCCCGAC 3607
Db 2941 CAGAGAGAAGTGTCTTGAAGACTACTTCCAGTGGTGTACTGTGTACCCCGAC 3000
Qy 3608 ACGGATGTGCAGCCCAAGTCTACGGCAAGCAGTGTGCAAGACTGCTCTAAGTCCA 3667
Db 3001 ACGGATGTGCAGCCCAAGTCTACGGCAAGCAGTGTGCTGCAAGACTGCTCTAAGTCCA 3060
Qy 3668 ACTTGTGA 3675
Db 3061 ACTTGTGA 3068

RESULT 14

BD170085
LOCUS BD170085 7668 bp DNA linear PAT 17-JAN-2003
DEFINITION Novel ADAMTS family polypeptide and gene encoding the same.
ACCESSION BD170085
VERSION BD170085.1 GI:27875897
KEYWORDS WO 0231163-A/7.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 7668)
AUTHORS Ohara,O., Nagase,T., Nomura,N., Yano,K., Murakami,K., Yasuda,S. and Kanazaki,K.
TITLE Novel ADAMTS family polypeptide and gene encoding the same
JOURNAL Patent: WO 0231163-A 7 18-APR-2002;
KAZUSA DNA RESEARCH INSTITUTE, MITSUBISHI PHARMA CORP. OSAMU OHARA,
KAZUHIRO NAGASE, NOBUO NOMURA, KAZUHIRO YANO, KOJI MURAKAMI, SHINICHIRO
YASUDA, KOJI KANZAKI
COMMENT OS Artificial Sequence
PN WO 0231163-A/7
PD 18-APR-2002
PF 11-OCT-2001 WO 2001JP008913

PR 11-OCT-2000 JP 00P 311309,02-APR-2001 JP 01P 102905 PI
OSAMU OHARA, TAKAHIRO NAGASE, NOBUO NOMURA, KAZUHIRO YANO, KOJI
MURAKAMI,
PI SHINICHIRO YASUDA, KOJI KANZAKI
PC C12N15/57, C12N9/64, C12N5/00, C12N1/19, C12N1/21, C12N1/15, C07K16/
40, C12Q1/68,
PC G01N33/50
CC Description of Artificial Sequence: base sequence of the CC
plasmid
CC pFastBacl-MS/Ht-PJ01256-1
FH Key Location/Qualifiers
FT source 1..7668
FT /organism='Artificial Sequence'.
FEATURES
source
ORIGIN
Query Match 77.5%; Score 2849.4; DB 6; Length 7668;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2850; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 822 CTTAGGGCATTAAGCGCTCTCTTCTGAGGTCCCATAGAAATGAAGAACTGAACCTGGAGAC 881
Db 4106 CTTAGGGCATTAAGCGCTCTCTTCTGAGGTCCCATAGAAATGAAGAACTGAACCTGGAGAC 4165
Qy 882 CTTGGTGTGTGTCACAAAAGATGATGCAAAACCATGGCCATGAAATATACACACCTA 941
Db 4166 CTTGGTGTGTGTCACAAAAGATGATGCAAAACCATGGCCATGAAATATACACACCTA 4225
Qy 942 CGTGTCAACGATCTCAACATGGTATCTGCTTTTATTCAAAGATGGAACAATAGGAGGAAA 1001
Db 4226 CGTGTCAACGATCTCAACATGGTATCTGCTTTTATTCAAAGATGGAACAATAGGAGGAAA 4285
Qy 1002 CATCAACATTTGCAATTTGATGCTGATCTTCTAGAGATGAACAGCCAGACTGGTGTAT 1061
Db 4286 CATCAACATTTGCAATTTGATGCTGATCTTCTAGAGATGAACAGCCAGACTGGTGTAT 4345
Qy 1062 AAGTCACACGACACACACCTTAAAGTAGCTTCTGCGAGTGGCAGTCTGATTTGATGGG 1121
Db 4346 AAGTCACACGACACACACCTTAAAGTAGCTTCTGCGAGTGGCAGTCTGATTTGATGGG 4405
Qy 1122 GAAAGATGGGACTCGTCAATGACACGCACTTTACTGACTGGTCTGATATATGTTCTGT 1181
Db 4406 GAAAGATGGGACTCGTCAATGACACGCACTTTACTGACTGGTCTGATATATGTTCTGT 4465
Qy 1182 GAAGAATGAGCCCTGTGACACTTTTGGGATTTGCAACCCATAGTGAATGTCTAGTAATA 1241
Db 4466 GAAGAATGAGCCCTGTGACACTTTTGGGATTTGCAACCCATAGTGAATGTCTAGTAATA 4525
Qy 1242 TCGCAGCTGCACGATTAATGAAGATACAGGTCTTTGGACTGGCCCTTACCATTGGCCCATGA 1301
Db 4526 TCGCAGCTGCACGATTAATGAAGATACAGGTCTTTGGACTGGCCCTTACCATTGGCCCATGA 4585
Qy 1302 GTCTGGACACAACTTTGGCATGATTCATGATGGAAGAGGGAACATGTGTAAAAAGTCCGA 1361
Db 4586 GTCTGGACACAACTTTGGCATGATTCATGATGGAAGAGGGAACATGTGTAAAAAGTCCGA 4645
Qy 1362 GGGCAACATCATGTCCCTACATTTGGCAGGACGCAATGGAGTCTTCTCCCTGGTCACCCCTG 1421
Db 4646 GGGCAACATCATGTCCCTACATTTGGCAGGACGCAATGGAGTCTTCTCCCTGGTCACCCCTG 4705
Qy 1422 CAGCCGCAAGTATCTACAAAATTTCTAAGCACCGCTCAAGCTATTCGCTTGTCTGATCA 1481
Db 4706 CAGCCGCAAGTATCTACAAAATTTCTAAGCACCGCTCAAGCTATTCGCTTGTCTGATCA 4765
Qy 1482 GCCAAAGCCTGTGAAGGAATACAGTATCTGAGAAATTCGCCAGGAATATATATGATGC 1541
Db 4766 GCCAAAGCCTGTGAAGGAATACAGTATCTGAGAAATTCGCCAGGAATATATATGATGC 4825
Qy 1542 AAACACACAGTGCAGAGTGGCAGTTTCGGAGAGAAAGCCAAAGCTCTGCTGCTGGACTTTAA 1601

LOCUS CQ805084 3287 bp DNA linear PAT 23-AUG-2004
DEFINITION Sequence 1053 from Patent EP1447413.
ACCESSION CQ805084
VERSION CQ805084.1 GI:51508796
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Isogai,T., Yamamoto,J., Nishikawa,T., Isono,Y., Sugiyama,T.,
Otsuki,T., Wakamatsu,A., Ishii,S., Nagai,K. and Irie,R.
TITLE Full-length human cDNA
JOURNAL Patent: EP 1447413-A 1053 18-AUG-2004;
Research Association for Biotechnology (JP)
FEATURES
Location/Qualifiers
1..3287
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 73.3%; Score 2692.8; DB 6; Length 3287;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGAAGCCCGCGCGCGATGGCGGGCTTGGCGCGCTGGGATGCTGCTGGCGCAG 60
Db 139 ATGAAGCCCGCGCGCGATGGCGGGCTTGGCGCGCTGGGATGCTGCTGGCGCAG 198
Qy 61 GTGGCCGAGGAGCAGCTGCGTGGCCATGGGACCCGACGCGCGCGCTGGGACCCG 120
Db 199 GTGGCCGAGGAGCAGCTGCGTGGCCATGGGACCCGACGCGCGCGCTGGGACCCG 258
Qy 121 AGCGTCCCGGCTCTCTCCACCGCGGAGCGCGCGCGCTGGATGCGTGGGAGCCCG 180
Db 259 AGCGTCCCGGCTCTCTCCACCGCGGAGCGCGCGCGCTGGATGCGTGGGAGCCGA 318
Qy 181 GACCTGGTCTCTGCTACGAGGTGACCAAGCGGCGATTAAGTGTCCCATGAATCATG 240
Db 319 GACCTGGTCTCTGCTACGAGGTGACCAAGCGGCGATTAAGTGTCCCATGAATCATG 378
Qy 241 CACCATCAGCGCGGAGAGAGCAGTGGCGGTGTCGAGGTGAGTCTCTTCACTTCGG 300
Db 379 CACCATCAGCGCGGAGAGAGCAGTGGCGGTGTCGAGGTGAGTCTCTTCACTTCGG 438
Qy 301 CTGAAAGGCTTCCAGGACGACTTCCACGTGGATCTCAGGACTTCCAGCAGCCTAGTGGCT 360
Db 439 CTGAAAGGCTTCCAGGACGACTTCCACGTGGATCTCAGGACTTCCAGCAGCCTAGTGGCT 498
Qy 361 CTTGGCTTTATTTGTGAG 420
Db 499 CTTGGCTTTATTTGTGAG 558
Qy 421 CCAGAGAGACTTCTGTTTCTATCAAGGCTCTTTGCGATACACAGAGAGAGAGAGAGAG 480
Db 559 CCAGAGAGACTTCTGTTTCTATCAAGGCTCTTTGCGATACACAGAGAGAGAGAGAGAG 618
Qy 481 CTTTCAACCTGCAAGGCTTGTGAGGATGATACGAAACAGAGAGAGAGAGAGAGAGAGAG 540
Db 619 CTTTCAACCTGCAAGGCTTGTGAGGATGATACGAAACAGAGAGAGAGAGAGAGAGAGAG 678
Qy 541 AGGCCACTTCTTCAACCTCTCATGGAACCTGCGGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 679 AGGCCACTTCTTCAACCTCTCATGGAACCTGCGGAGAGAGAGAGAGAGAGAGAGAGAG 738
Qy 601 TCCAGAGTACTGTACAAGAGATCCACAGAGCCCATGCTCTCGGCGCCAGTGGAGTCTG 660
Db 739 TCCAGAGTACTGTACAAGAGATCCACAGAGCCCATGCTCTCGGCGCCAGTGGAGTCTG 798
Qy 661 GTGACCTCAAGAGATGGAGGTGGGACATCAACCCCTGCAAGAGAGAGAGAGAGAGAG 720
Db 799 GTGACCTCAAGAGATGGAGGTGGGACATCAACCCCTGCAAGAGAGAGAGAGAGAGAG 858

Qy 721 GGAATGCCCAAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCGACCTCCC 780
Db 859 GGAATGCCCAAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCGACCTCCC 918
Qy 781 AAGGAAGACTCTTCTATCTTGCAGATGATTAAGTCTTGTCTTACGGCATTAAGCGCTCT 840
Db 919 AAGGAAGACTCTTCTATCTTGCAGATGATTAAGTCTTGTCTTACGGCATTAAGCGCTCT 978
Qy 841 CTTCTGAGGTCCCATGAATAAGAACTGAACCTGGAGACCTTGGTGGTGGTGGCAAA 900
Db 979 CTTCTGAGGTCCCATGAATAAGAACTGAACCTGGAGACCTTGGTGGTGGTGGCAAA 1038
Qy 901 AAGATGATGCAAAACCATGGCCATGAAATATCACCACTACGTGCTCACGATACTCAAC 960
Db 1039 AAGATGATGCAAAACCATGGCCATGAAATATCACCACTACGTGCTCACGATACTCAAC 1098
Qy 961 ATGGTATCTGCTTTATTCAAAGATGGAACATAGAGGAAACATCAACATTTGCAATTGTA 1020
Db 1099 ATGGTATCTGCTTTATTCAAAGATGGAACATAGAGGAAACATCAACATTTGCAATTGTA 1158
Qy 1021 GGTCTGATTTCTTAGAAGATGAACAGCAGGACTGGTGATAAAGTCACCAACGACCAAC 1080
Db 1159 GGTCTGATTTCTTAGAAGATGAACAGCAGGCTGGTGATAAAGTCACCAACGACCAAC 1218
Qy 1081 ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGATTTGATGGGAAAGATGGGACTCGTCA 1140
Db 1219 ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGATTTGATGGGAAAGATGGGACTCGTCA 1278
Qy 1141 GACCACGCGCATCTTACTGACTGGTCTGGATATATGTTCTCTGGAGAATGAGCCCTGTGAC 1200
Db 1279 GACCACGCGCATCTTACTGACTGGTCTGGATATATGTTCTCTGGAGAATGAGCCCTGTGAC 1338
Qy 1201 ACTTTGGGATTTGCAACCCATTAAGTGGAAATGTGTAGTAAATATCGCAGCTGACCAATTA 1260
Db 1339 ACTTTGGGATTTGCAACCCATTAAGTGGAAATGTGTAGTAAATATCGCAGCTGACCAATTA 1398
Qy 1261 GAAGATACAGTCTTGGACTGGCCTTCCCAATGGCCCATGAGCTGTGGACACAACTTTGGC 1320
Db 1399 GAAGATACAGTCTTGGACTGGCCTTCCCAATGGCCCATGAGCTGTGGACACAACTTTGGC 1458
Qy 1321 ATGATTCTAAGCAGCGCTCAAGCTATCTGCTTCTGCTGATCCAGCGCCAGTATCTACAC 1380
Db 1459 ATGATTCTAAGCAGCGCTCAAGCTATCTGCTTCTGCTGATCCAGCGCCAGTATCTACAC 1518
Qy 1381 ACATTTGGCAGGACGCAATGGAGTCTTCTCTTGGTCAACCTGAGCGCCAGTATCTACAC 1440
Db 1519 ACATTTGGCAGGACGCAATGGAGTCTTCTCTTGGTCAACCTGAGCGCCAGTATCTACAC 1578
Qy 1441 AAATTTCTAAGCAGCGCTCAAGCTATCTGCTTCTGCTGATCCAGCGCCAGTATCTACAC 1500
Db 1579 AAATTTCTAAGCAGCGCTCAAGCTATCTGCTTCTGCTGATCCAGCGCCAGTATCTACAC 1638
Qy 1501 TACAAGTATCTGAGAAATTTGCCAGGAGAAATATATGATGCAAAACACAGTGAAGTGG 1560
Db 1639 TACAAGTATCTGAGAAATTTGCCAGGAGAAATATATGATGCAAAACACAGTGAAGTGG 1698
Qy 1561 CAGTTTCGAGAGAAAGCCAAAGCTCTGCATCTGGAACCTTTAAAAGGACATCTGTAAGGCC 1620
Db 1699 CAGTTTCGAGAGAAAGCCAAAGCTCTGCATCTGGAACCTTTAAAAGGACATCTGTAAGGCC 1758
Qy 1621 CTGTGGTGCCATCGTATTTGGAAGGAAATGTGAGACTAAATTTATGCCAGCAGCAGAAAGGC 1680
Db 1759 CTGTGGTGCCATCGTATTTGGAAGGAAATGTGAGACTAAATTTATGCCAGCAGCAGAAAGGC 1818
Qy 1681 ACATTTTGTGGGATGATGATGTTGGTCCCGGGAGGACAGTGTGTGAATATGTGTGATGAA 1740
Db 1819 ACATTTTGTGGGATGATGATGTTGGTCCCGGGAGGACAGTGTGTGAATATGTGTGATGAA 1878
Qy 1741 GGCCCCAAGCCCAACCCATGGCCACTGGTCCGAGCTGGTCTTCTTGGTCCCAATGCTCCAGG 1800
Db 1879 GGCCCCAAGCCCAACCCATGGCCACTGGTCCGAGCTGGTCTTCTTGGTCCCAATGCTCCAGG 1938

```
QY 1801 ACCTGGGAGGGGAGTATCTCATAGAGTCGGCTCTGCACCAACCCCAAGCCATCGCAT 1860
Db 1939 ACCTGGGAGGGGAGTATCTCATAGAGTCGGCTCTGCACCAACCCCAAGCCATCGCAT 1998
QY 1861 GGAGGCAAGTTCCTGAGGGCTCCACTCGCACTCTGAAGCTCTGCAACAGTCAAGAAATGT 1920
Db 1999 GGAGGCAAGTTCCTGAGGGCTCCACTCGCACTCTGAAGCTCTGCAACAGTCAAGAAATGT 2058
QY 1921 CCCCAGGACAGTCTTGAATTCCTGCTCGCTAGTGTCCGAGCAACACAGCAGCAGATTC 1980
Db 2059 CCCCAGGACAGTCTTGAATTCCTGCTCGCTAGTGTCCGAGCAACACAGCAGCAGATTC 2118
QY 1981 AGAGGGGGGCACTACAGTGGAGCCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAA 2040
Db 2119 AGAGGGGGGCACTACAGTGGAGCCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAA 2178
QY 2041 CTCTACTGTATCCGACAGGATTTGATTTCTTTCTTTTGTCAATAAAGTCAAGAT 2100
Db 2179 CTCTACTGTATCCGACAGGATTTGATTTCTTTCTTTTGTCAATAAAGTCAAGAT 2238
QY 2101 GGGACTCCATGTCTGGAGGATAGCCGTAATGTTTGTATAGATGGGATATGTGAGAGATT 2160
Db 2239 GGGACTCCATGTCTGGAGGATAGCCGTAATGTTTGTATAGATGGGATATGTGAGAGATT 2298
QY 2161 GGNATGTGACAAATGTCCTTGGATCTGATCTGTTGAAAGACGTCTGTGGGGTGTAAACGGG 2220
Db 2299 GGNATGTGACAAATGTCCTTGGATCTGATCTGTTGAAAGACGTCTGTGGGGTGTAAACGGG 2358
QY 2221 AATAACTCAGCCTGCAGATTACAGGGGTCTCTACCAAGCACCAACACCAACCAACCGAG 2280
Db 2359 AATAACTCAGCCTGCAGATTACAGGGGTCTCTACCAAGCACCAACACCAACCAACCGAG 2418
QY 2281 TATTTATCATGTGTCAACATTCCTTCTGGAGCCCGAGTATCGCATCTATGAAATGAAC 2340
Db 2419 TATTTATCATGTGTCAACATTCCTTCTGGAGCCCGAGTATCGCATCTATGAAATGAAC 2478
QY 2341 GTCTCTACCTCTTACATTTCTGTGGGGAATGCCCTCAGAAAGTACTTACCTGAAATGGGCAC 2400
Db 2479 GTCTCTACCTCTTACATTTCTGTGGGGAATGCCCTCAGAAAGTACTTACCTGAAATGGGCAC 2538
QY 2401 TGGACCGTGGACTGGCCCGGGGTACAAATTTTCGGGCACACTTTTCGACTACAGACGG 2460
Db 2539 TGGACCGTGGACTGGCCCGGGGTACAAATTTTCGGGCACACTTTTCGACTACAGACGG 2598
QY 2461 TCCTATAATGAGCCCGAGAACTTAAATCGCTACTGGACCAACCAACGAGACACTGATTGTG 2520
Db 2599 TCCTATAATGAGCCCGAGAACTTAAATCGCTACTGGACCAACCAACGAGACACTGATTGTG 2658
QY 2521 GAGCTGCTGTTTCAGGGAAGGAACCCGGGTGTGCTTGGGAATATCTCCATGCCCTGCTTG 2580
Db 2659 GAGCTGCTGTTTCAGGGAAGGAACCCGGGTGTGCTTGGGAATATCTCCATGCCCTGCTTG 2718
QY 2581 GGGACCGAGAGCAGCCCTGCGCCAGCCAGCTACACTTGGGCCATCGTGGCTCTGAG 2640
Db 2719 GGGACCGAGAGCAGCCCTGCGCCAGCCAGCTACACTTGGGCCATCGTGGCTCTGAG 2778
QY 2641 TGCCTCGTGTCTCTGGAGGGGACAGATGACCGTGAAGAGGGGCTGTACAGAGA 2696
Db 2779 TGCCTCGTGTCTCTGGAGGGGACAGATGACCGTGAAGAGGGGCTGTACAGAGA 2834
```

Search completed: June 13, 2005, 20:47:01
Job time : 15143 secs

THIS PAGE BLANK (USPTO)

Result No.	Query			Length	DB	ID	Description
	Score	Match	%				
1	3675	100.0	3675	9	ADA50761		Ada50761 cDNA enco
2	3675	100.0	3675	10	ABSS57767		Abss57767 Novel hum
3	3675	100.0	4042	9	ADA50763		Ada50763 cDNA enco
4	3675	100.0	4042	10	ABSS57768		Abss57768 Novel hum
5	3673.4	100.0	3675	6	ALA43654		Aal43654 Human MDT
6	3673.4	100.0	3675	12	ACC95689		Acc95689 Human pro
7	3671.8	99.9	3675	6	AA\$971174		Aas971174 Human met
8	3671.8	99.9	3675	9	ACF57047		Acf57047 Human ADA
9	3671.8	99.9	5610	6	ABK49822		Abk49822 Human CDN
10	3668.8	99.8	8435	6	ABK49825		Abk49825 Plasmid p
11	3668.8	99.8	8505	6	ABK49826		Abk49826 Plasmid p
12	3182.8	86.6	3227	10	ACCO0343		Acc00343 Human ADA
13	3169.2	86.2	4902	13	ACN43153		Actn43153 Human dia
14	3135.2	85.3	4888	6	ABK86137		Abk86137 cDNA enco
15	3064.8	83.4	4234	6	ABK49821		Abk49821 Human CDN
16	2849.4	77.5	7668	6	ABK49827		Abk49827 Plasmid p
17	2692.8	73.3	3287	13	ADRO7547		Adro7547 Full lengt
18	2055.6	55.9	2997	6	ABS59323		Abes59323 Human zin
19	2055.6	55.9	2997	10	ADJ38419		Adj38419 Human CDN
20	2022.8	55.0	2902	6	ABS59325		Abes59325 Human ADA

PR 16-AUG-2000; 2000US-0225852P.
 PA (LEXI-) LEXICON GENETICS INC.
 XX Friddle CJ, Hilbun E;
 PI WPI; 2003-074103/07.
 DR P-PSDB; ABG72431.
 XX
 DR New nucleic acid encoding novel human protein (NHP), useful for the
 PT preparation of a medicament for diagnosing, preventing or treating
 PT disorders or diseases associated with the NHP, e.g., obesity.
 XX
 PS Claim 2; Col 17-20; 17pp; English.
 XX
 CC The invention describes a new isolated nucleic acid comprising a sequence
 CC that encodes a fully defined protein sequence comprising 491 or 1224
 CC amino acids or that hybridises under stringent conditions with the 3675-
 CC bp sequence or its complement. The nucleic acid is useful for the
 CC preparation of a medicament for diagnosing, preventing or treating
 CC disorders (e.g. using gene therapy) or diseases associated with the novel
 CC human proteins, e.g., obesity. This sequence encodes a novel human
 CC protease
 XX
 SQ Sequence 3675 BP; 873 A; 1008 C; 1020 G; 774 T; 0 U; 0 Other;
 Query Match 100.0%; Score 3675; DB 10; Length 3675;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3675; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGCCCGCGCGCGGATGCGGGCTTGGCGCGCTGTGGATGCTGTGGCGCAG 60
 DB |||||
 QY 1 ATGAAGCCCGCGCGCGGATGCGGGGCTTGGCGCGCTGTGGATGCTGTGGCGCAG 60
 DB |||||
 QY 61 GTGCGCCAGCAGGACCTGCGTGGCGCATGGGACCGCGCAGCGCGCGCGCGCGCG 120
 DB |||||
 QY 61 GTGCGCCAGCAGGACCTGCGTGGCGCATGGGACCGCGCAGCGCGCGCGCGCGCG 120
 DB |||||
 QY 121 AGCGTCCCGCGTCTCTCTCCACCGCGGAGCGCGCGGCTGGATGGAAGCGCGAAT 180
 DB |||||
 QY 121 AGCGTCCCGCGTCTCTCTCCACCGCGGAGCGCGCGGCTGGATGGAAGCGCGAAT 180
 DB |||||
 QY 181 GACCTGTCTCTGCTACGAGGTGACCAAGGGCGGATTA CGTGTCCATGAATCATG 240
 DB |||||
 QY 181 GACCTGTCTCTGCTACGAGGTGACCAAGGGCGGATTA CGTGTCCATGAATCATG 240
 DB |||||
 QY 241 CACCATCAGCGCGGAGAGCAGTGGCGGTGTCGAGGTTGAGTCTCTT CACCTTCGG 300
 DB |||||
 QY 241 CACCATCAGCGCGGAGAGCAGTGGCGGTGTCGAGGTTGAGTCTCTT CACCTTCGG 300
 DB |||||
 QY 301 CTGAAAGGCTCCAGGACGACTTCCACGTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360
 DB |||||
 QY 301 CTGAAAGGCTCCAGGACGACTTCCACGTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360
 DB |||||
 QY 361 CTGGGCTTTATGTGACAGAGTGGGAAAGACAGGACCTAAGTCTCTT CACCTTCGG 420
 DB |||||
 QY 361 CTGGGCTTTATGTGACAGAGTGGGAAAGACAGGACCTAAGTCTCTT CACCTTCGG 420
 DB |||||
 QY 421 CCAGAGGACTTCTGTTTCTATCAAGGCTTTTGCATCACACAGAACTCTCTCAGTGGCC 480
 DB |||||
 QY 421 CCAGAGGACTTCTGTTTCTATCAAGGCTTTTGCATCACACAGAACTCTCTCAGTGGCC 480
 DB |||||
 QY 481 CTTTCAACCTGCAAGGCTTGTGAGGATGATACGAAACAGAGAGGCGGATTA TCTCTTA 540
 DB |||||
 QY 481 CTTTCAACCTGCAAGGCTTGTGAGGATGATACGAAACAGAGAGGCGGATTA TCTCTTA 540
 DB |||||
 QY 541 AGGCCACTTCTTCA CACTCTCATGGAACTCGGACAGCTGCCAAGGCGAGCTCGCCA 600
 DB |||||
 QY 541 AGGCCACTTCTTCA CACTCTCATGGAACTCGGACAGCTGCCAAGGCGAGCTCGCCA 600
 DB |||||
 QY 601 TCCACAGTACTGTACAAGAGATCCACAGGCCCACTGCTCTCTGGGCGAGTGGCTCTG 660
 DB |||||
 QY 601 TCCACAGTACTGTACAAGAGATCCACAGGCCCACTGCTCTCTGGGCGAGTGGCTCTG 660
 DB |||||

QY 661 GTGACCTCAAGGACATGGGAGCTGGCA CATCAACCCCTGACAGCAGCAGCTTGCCTG 720
 DB |||||
 QY 661 GTGACCTCAAGGACATGGGAGCTGGCA CATCAACCCCTGACAGCAGCAGCTTGCCTG 720
 DB |||||
 QY 721 GGACTGCCACAAAGCAGCATTTCTGTGGAAGCAGCAAGAAATACATGCCCCAGCCTCC 780
 DB |||||
 QY 721 GGACTGCCACAAAGCAGCATTTCTGTGGAAGCAGCAAGAAATACATGCCCCAGCCTCC 780
 DB |||||
 QY 781 AAGGAAGACCTCTTCATCTTTCAGATGAGTATAAGTCTTGTCTTACGCGATAGCGCTCT 840
 DB |||||
 QY 781 AAGGAAGACCTCTTCATCTTTCAGATGAGTATAAGTCTTGTCTTACGCGATAGCGCTCT 840
 DB |||||
 QY 841 CTTCTGAGTCCCATAGAAATGAAGACTGAACGTGGAGACCTTGGTGGTGTGCACAAA 900
 DB |||||
 QY 841 CTTCTGAGTCCCATAGAAATGAAGACTGAACGTGGAGACCTTGGTGGTGTGCACAAA 900
 DB |||||
 QY 901 AAGATGATGCAAAACCATGGCCATGAAATATCACCACTACGTGCTCAGCATCTCAAC 960
 DB |||||
 QY 901 AAGATGATGCAAAACCATGGCCATGAAATATCACCACTACGTGCTCAGCATCTCAAC 960
 DB |||||
 QY 961 ATGGTATCTCTCTTATTCAAGATGGAACATAGAGGAAACATCAACATTCGAATTGTA 1020
 DB |||||
 QY 961 ATGGTATCTCTCTTATTCAAGATGGAACATAGAGGAAACATCAACATTCGAATTGTA 1020
 DB |||||
 QY 1021 GGTCTGATTTCTTAGAAGATGAACAGCAGGACTGGGTAGTAAAGTACACGAGCAGAC 1080
 DB |||||
 QY 1021 GGTCTGATTTCTTAGAAGATGAACAGCAGGACTGGGTAGTAAAGTACACGAGCAGAC 1080
 DB |||||
 QY 1081 ACCTTAAGTAGCTTCTGCGAGTGGCAGTCTGGATGATGGGAAAGATGGGACTCGT CAT 1140
 DB |||||
 QY 1081 ACCTTAAGTAGCTTCTGCGAGTGGCAGTCTGGATGATGGGAAAGATGGGACTCGT CAT 1140
 DB |||||
 QY 1141 GACCACGCGCATTTACTGACTGGTCTGGATATATGTTCTGGAGAGATAGCCCTGTGAC 1200
 DB |||||
 QY 1141 GACCACGCGCATTTACTGACTGGTCTGGATATATGTTCTGGAGAGATAGCCCTGTGAC 1200
 DB |||||
 QY 1201 ACTTTGGGATTTGCCCATTAAGTGAATGTAGTAAATATCGCAGCTGCAGATTAAT 1260
 DB |||||
 QY 1201 ACTTTGGGATTTGCCCATTAAGTGAATGTAGTAAATATCGCAGCTGCAGATTAAT 1260
 DB |||||
 QY 1261 GAAGATACAGTCTTTGAGCTGGCTTCCACCATTTGCCATGAGCTGGACACAATTTGGC 1320
 DB |||||
 QY 1261 GAAGATACAGTCTTTGAGCTGGCTTCCACCATTTGCCATGAGCTGGACACAATTTGGC 1320
 DB |||||
 QY 1321 ATGATTATGATGAGAGAGGAA CATGTGTAAAGTCCGAGGGCAACATCATGTCCCCT 1380
 DB |||||
 QY 1321 ATGATTATGATGAGAGAGGAA CATGTGTAAAGTCCGAGGGCAACATCATGTCCCCT 1380
 DB |||||
 QY 1381 ACATTGCGAGCGCAATGGAGTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 DB |||||
 QY 1381 ACATTGCGAGCGCAATGGAGTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 DB |||||
 QY 1441 AAATTTCTAAGCACCGCTCAAGCTATCTGCTTGTGATCAGCAGCAAGCTGTGAAGGAA 1500
 DB |||||
 QY 1441 AAATTTCTAAGCACCGCTCAAGCTATCTGCTTGTGATCAGCAGCAAGCTGTGAAGGAA 1500
 DB |||||
 QY 1501 TACAAGTATCTGAGAAATTTGCCAGAGAAATTAATGATGCAAAACACACAGTGAAGTGG 1560
 DB |||||
 QY 1501 TACAAGTATCTGAGAAATTTGCCAGAGAAATTAATGATGCAAAACACACAGTGAAGTGG 1560
 DB |||||
 QY 1561 CAGTTCCGAGAGAAAGCCCAAGCTCTGCACTGTGGACTTTTAAAGAGACATCTGTAAAGCC 1620
 DB |||||
 QY 1561 CAGTTCCGAGAGAAAGCCCAAGCTCTGCACTGTGGACTTTTAAAGAGACATCTGTAAAGCC 1620
 DB |||||
 QY 1621 CTGTGTGCGCATCTGATTTGGAAGGAAATGTGAGACTAAATTTATGCCAGCAGCAGAGGC 1680
 DB |||||
 QY 1621 CTGTGTGCGCATCTGATTTGGAAGGAAATGTGAGACTAAATTTATGCCAGCAGCAGAGGC 1680
 DB |||||
 QY 1681 ACAATTTGTGGCGATGACATGTGGTGGCGGAGGACAGTGTGTGAAATATGGTATGAA 1740
 DB |||||
 QY 1681 ACAATTTGTGGCGATGACATGTGGTGGCGGAGGACAGTGTGTGAAATATGGTATGAA 1740
 DB |||||

Qy	1741	GGCCCCAAGCCACCCATGGCCGACATGGTCTTCTTGGTCCCAAGTCTCCAGG	1800
Db	1741	GGCCCCAAGCCACCCATGGCCACATGGCCACTGGTGGAGCTGGTCTTCTTGGTCCCAATGTCTCCAGG	1800
Qy	1801	ACCTCGCGAGGGGAGTATCTCATAGGAGTGCCTCTGTGCAACCAACCCAGCCCATCCGAT	1860
Db	1801	ACCTCGCGAGGGGAGTATCTCATAGGAGTGCCTCTGTGCAACCAACCCAGCCCATCCGAT	1860
Qy	1861	GGAGGGAAGTTCTGTGAGGGCTCCACTCGCACTCTGAAGCTCTGCAACAGTCAAGAAATGT	1920
Db	1861	GGAGGGAAGTTCTGTGAGGGCTCCACTCGCACTCTGAAGCTCTGCAACAGTCAAGAAATGT	1920
Qy	1921	CCCCGGGACAGTGTGACTTCGGTGTCTGCTCAGTGTGCCGAGCAACAGCAGACGATTC	1980
Db	1921	CCCCGGGACAGTGTGACTTCGGTGTCTGCTCAGTGTGCCGAGCAACAGCAGACGATTC	1980
Qy	1981	AGAGGGCGGCACTCAAGTGGAGCCCTTACACTCAAGTGAAGAGATCAGGACTTATGCAAA	2040
Db	1981	AGAGGGCGGCACTCAAGTGGAGCCCTTACACTCAAGTGAAGAGATCAGGACTTATGCAAA	2040
Qy	2041	CTCTACTGTATCGCAGAAAGGATTTGATTTCTTTCTTTTGTCAAAATAAAGTCAAAAGAT	2100
Db	2041	CTCTACTGTATCGCAGAAAGGATTTGATTTCTTTCTTTTGTCAAAATAAAGTCAAAAGAT	2100
Qy	2101	GGGACTCCATGCTCGAGGATAGCCGTAATGTTGTATAGATGGGATATGTGAGAGATTT	2160
Db	2101	GGGACTCCATGCTCGAGGATAGCCGTAATGTTGTATAGATGGGATATGTGAGAGATTT	2160
Qy	2161	GGATGTGACAAATGCTTGGATCTGAATGCTTGGATGCTGTTGAAGCGTCTGTGGGGTGTGTAACGGG	2220
Db	2161	GGATGTGACAAATGCTTGGATCTGAATGCTTGGATGCTGTTGAAGCGTCTGTGGGGTGTGTAACGGG	2220
Qy	2221	AATAAATCAGGCTGCGACGATTTACAGGGGTTCTTACACCAAGCACACCAACCAACGAG	2280
Db	2221	AATAAATCAGGCTGCGACGATTTACAGGGGTTCTTACACCAAGCACACCAACCAACGAG	2280
Qy	2281	TATTATCAATGGTCAACATTTCTTCTGGAGCCCGGAGTATCCGACTCTATGAATGAAC	2340
Db	2281	TATTATCAATGGTCAACATTTCTTCTGGAGCCCGGAGTATCCGACTCTATGAATGAAC	2340
Qy	2341	GTCTCTACCTCTACATTTCTGTGGCAATGCCCTCAGAGAGTACTACCTGGAATGGGCAC	2400
Db	2341	GTCTCTACCTCTACATTTCTGTGGCAATGCCCTCAGAGAGTACTACCTGGAATGGGCAC	2400
Qy	2401	TGGACCGTGGACTGCGCCCGGCGGTACAAAATTTTTCGGGCACACTCTTTCGACTACAGACGG	2460
Db	2401	TGGACCGTGGACTGCGCCCGGCGGTACAAAATTTTTCGGGCACACTCTTTCGACTACAGACGG	2460
Qy	2461	TCCTATAATAGCCCGAGAACTTAATCGCTACTGGACCAACCAACGAGACACTGATTTGTG	2520
Db	2461	TCCTATAATAGCCCGAGAACTTAATCGCTACTGGACCAACCAACGAGACACTGATTTGTG	2520
Qy	2521	GAGCTGCTGTTTTCAGGNAAGAAACCGGGTGTTCCTTGGGNAATCTCCATGCCCTCGCTTG	2580
Db	2521	GAGCTGCTGTTTTCAGGNAAGAAACCGGGTGTTCCTTGGGNAATCTCCATGCCCTCGCTTG	2580
Qy	2581	GGGACCGAGAAAGCAGCCCTGCGCCAGCCAGCTACACTTGGGGCCATCGTGGCTCTGAG	2640
Db	2581	GGGACCGAGAAAGCAGCCCTGCGCCAGCCAGCTACACTTGGGGCCATCGTGGCTCTGAG	2640
Qy	2641	TGCTCCGTGTCTCGGAGGGGGACAGATGACCGTGAAGAGGGGCTGCTACAGAGACCTG	2700
Db	2641	TGCTCCGTGTCTCGGAGGGGGACAGATGACCGTGAAGAGGGGCTGCTACAGAGACCTG	2700
Qy	2701	AAGTTTCAAGTAAATATGTCTTCTGCAATCCCAAGACACGACTGTGTCAGGGGCTGGTG	2760
Db	2701	AAGTTTCAAGTAAATATGTCTTCTGCAATCCCAAGACACGACTGTGTCAGGGGCTGGTG	2760
Qy	2761	CTTTGCAAGTATCTGCTCTGTCTCCAGAGTGGTCCGTGGGAACTCGAGTGCCTGCAGT	2820
Db	2761	CTTTGCAAGTATCTGCTCTGTCTCCAGAGTGGTCCGTGGGAACTCGAGTGCCTGCAGT	2820
Qy	2821	CGGACGTGTGCGGGGGTGGCCAGAGCGCCCGTGCAGTGCACACAGCGGGGTGCATAT	2880

Db	2821	CGGACGTGTGGCGGGGTGGCCAGAGCCGCCCGGTGCAAGTGCACACGGCGGGTGCACAT	2880
Qy	2881	GACTCGGAGCCAGTCCGCGCCAGCCTGTGCCCCTCAGCTGCTCCCTCCAGCAGCGCAGGCC	2940
Db	2881	GACTCGGAGCCAGTCCGCGCCAGCCTGTGCCCCTCAGCTGCTCCCTCCAGCAGCGCAGGCC	2940
Qy	2941	TGCAACTCTCAGAGCTGCCCAACCTGCATGAGAGCGCCGCGGCCCTTGGGAGAGTGTCTCACAC	3000
Db	2941	TGCAACTCTCAGAGCTGCCCAACCTGCATGAGAGCGCCGCGGCCCTTGGGAGAGTGTCTCACAC	3000
Qy	3001	ACCTGTGGGAAGGGGTGGAGGAACGGCGGCAAGTGGCTGTGAAGACACAAACCCCTCGGCC	3060
Db	3001	ACCTGTGGGAAGGGGTGGAGGAACGGCGGCAAGTGGCTGTGAAGACACAAACCCCTCGGCC	3060
Qy	3061	AGCGCGAGCTGTCTCCCGACGCTGTCTGCACCTCCGAGGCCAAGGCCAGAGTGCATGAA	3120
Db	3061	AGCGCGAGCTGTCTCCCGACGCTGTCTGCACCTCCGAGGCCAAGGCCAGAGTGCATGAA	3120
Qy	3121	GCCTGTCTGTCTCAGCGCTGCCCAAGCCCAAGAAGCTGCAGTGGTGTGTCCGCTGG	3180
Db	3121	GCCTGTCTGTCTCAGCGCTGCCCAAGCCCAAGAAGCTGCAGTGGTGTGTCCGCTGG	3180
Qy	3181	TCCAGTGTCTGTGACATGTGAAGAGGAACAACAGAAAGATTCTTAAATGTGCTGAA	3240
Db	3181	TCCAGTGTCTGTGACATGTGAAGAGGAACAACAGAAAGATTCTTAAATGTGCTGAA	3240
Qy	3241	AAGTATGTCTTGAAAGTATCGAGAGCTGGCCCTCAAAGAAGTGTCCACATTTGCCGAAG	3300
Db	3241	AAGTATGTCTTGAAAGTATCGAGAGCTGGCCCTCAAAGAAGTGTCCACATTTGCCGAAG	3300
Qy	3301	CCCAGCCTGGAGCTGGAAACGTGTGCTGCGCCCCGCTTCCATGCCCCAGGCACCCCCCATTT	3360
Db	3301	CCCAGCCTGGAGCTGGAAACGTGTGCTGCGCCCCGCTTCCATGCCCCAGGCACCCCCCATTT	3360
Qy	3361	GCTGCTGGGGAACCTCGAGGGGAGCTGTTTGGCTTCCCTCAGCTGGTCTCAGTGCACGGCC	3420
Db	3361	GCTGCTGGGGAACCTCGAGGGGAGCTGTTTGGCTTCCCTCAGCTGGTCTCAGTGCACGGCC	3420
Qy	3421	AGCTGTGGGGAGGCGTTACAGACGAGTCCGTGCAGTGCCTGGCTGGGGGCGCGCCGGCC	3480
Db	3421	AGCTGTGGGGAGGCGTTACAGACGAGTCCGTGCAGTGCCTGGCTGGGGGCGCGCCGGCC	3480
Qy	3481	TCAGGCTGCCCTCTCTGCACCAAGAGCCCTTCGGCTCCCTCGCTGGCTGCACACTCACTTCTGC	3540
Db	3481	TCAGGCTGCCCTCTCTGCACCAAGAGCCCTTCGGCTCCCTCGCTGGCTGCACACTCACTTCTGC	3540
Qy	3541	CCCATTCGAGAGAAGAAAGATGCTTGTGCAAGAATCTTCCACTGGTGTACTCTGGTA	3600
Db	3541	CCCATTCGAGAGAAGAAAGATGCTTGTGCAAGAATCTTCCACTGGTGTACTCTGGTA	3600
Qy	3601	CCCAGCAGCGGATGTGCAGGCCACAGTTCAGGCAAGCAGTGTGCAAGACTTGCTCT	3660
Db	3601	CCCAGCAGCGGATGTGCAGGCCACAGTTCAGGCAAGCAGTGTGCAAGACTTGCTCT	3660
Qy	3661	AAGTCCCAACTTGTGA	3675
Db	3661	AAGTCCCAACTTGTGA	3675
RESULT 3			
ID	ADA50763	standard; DNA; 4042 BP.	
XX	ADA50763;		
AC	ADA50763;		
DT	20-NOV-2003	(first entry)	
XX	Novel human protein (NHP) ORF and flanking sequence.		
DE	Novel human protein; NHP; protease; zinc metalloproteinase; gene therapy;		
KW	human; gene; ds.		
XX			

[illegible]

ABSS7768
ID ABS57768 standard; cDNA; 4042 BP.
XX AC
XX ABS57768;
XX DT
XX 05-FEB-2003 (first entry)
XX DE
XX cDNA encoding novel human protease #2.
XX KW
XX Human; protease; gene therapy; obesity; gene; ss.
XX OS
XX Homo sapiens.
XX FH
XX Key Location/Qualifiers
XX CDS 99..3773
XX FT /*tag= a
XX FT /product= "Novel human protease"
XX PN
XX US6448388-B1.
XX PD
XX 10-SEP-2002.
XX PF
XX 15-AUG-2001; 2001US-00930872.
XX PR
XX 16-AUG-2000; 2000US-0225952P.
XX XX
XX (LEXI-) LEXICON GENETICS INC.
XX PA
XX Friddle CJ, Hilbun E;
XX PI
XX WPI; 2003-074103/07.
XX DR
XX P-PSDB; ABG72431.
XX
XX New nucleic acid encoding novel human protein (NHP), useful for the
XX PT preparation of a medicament for diagnosing, preventing or treating
XX PT disorders or diseases associated with the NHP, e.g., obesity.
XX
XX Disclosure; Col 25-30; 17pp; English.
XX
XX The invention describes a new isolated nucleic acid comprising a sequence
XX CC that encodes a fully defined protein sequence comprising 491 or 1224
XX CC amino acids or that hybridises under stringent conditions with the 3675-
XX CC bp sequence or its complement. The nucleic acid is useful for the
XX CC preparation of a medicament for diagnosing, preventing or treating
XX CC disorders (e.g. using gene therapy) or diseases associated with the novel
XX CC human proteins, e.g., obesity. This sequence encodes a novel human
XX CC protease
XX
XX Sequence 4042 BP; 947 A; 1117 C; 1130 G; 848 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 3675; DB 10; Length 4042;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3675; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAGCCCGCGCGCGATGGCGGCTTGGCGGCGCTGTGATGCTGTCGCGCAG 60
DB |||||
99 ATGAAGCCCGCGCGCGATGGCGGCTTGGCGGCGCTGTGATGCTGTCGCGCAG 158
QY 61 GTGGCGGAGCAGCAGCTCGCTGCGGCATGGGACCCGACGCGGCGCTGGAGCCCG 120
DB 159 GTGGCGGAGCAGCAGCTCGCTGCGGCATGGGACCCGACGCGGCGCTGGAGCCCG 218
QY 121 AGCGTCCGCGTCTCTCCACCGGAGCGGCGGCTGGATGGAAGGGCGGAATAT 180
DB |||||
219 AGCGTCCGCGTCTCTCCACCGGAGCGGCGGCTGGATGGAAGGGCGGAATAT 278
QY 181 GACCTGGTCTCTGCTACGAGTTGACCA CAGGGCGGATTACGTCTCCCATGAATCATG 240
DB 279 GACCTGGTCTCTGCTACGAGTTGACCA CAGGGCGGATTACGTCTCCCATGAATCATG 338
QY 241 CACCATCAGCGCGGAGAGACAGTGGCCGTGTCGAGGTTGAGTCTCTTACCTTCGG 300
DB 339 CACCATCAGCGCGGAGAGACAGTGGCCGTGTCGAGGTTGAGTCTCTTACCTTCGG 398

QY 301 CTGAAGGCTCCAGCAGCACTTCCACGTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360
DB |||||
399 CTGAAGGCTCCAGCAGCACTTCCACGTGGATCTGAGGACTTCCAGCAGCCTAGTGGCT 458
QY 361 CTTGGCTTTATTTGTGAGACGTTGGGAAAGACAGGCACTAAAGTCTGTGTCAGACTTTACCG 420
DB 459 CTTGGCTTTATTTGTGAGACGTTGGGAAAGACAGGCACTAAAGTCTGTGTCAGACTTTACCG 518
QY 421 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGATCACACAGAAATCTCTCAGTGGCC 480
DB 519 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGATCACACAGAAATCTCTCAGTGGCC 578
QY 481 CTTTCAACCTGCCAAGGCTTGTCTAGGCATGATAGAAACAGAGAGGCGAGATTACTTCTTA 540
DB 579 CTTTCAACCTGCCAAGGCTTGTCTAGGCATGATAGAAACAGAGAGGCGAGATTACTTCTTA 638
QY 541 AGGCCACTTCTTTCACACCTCTCATGAACTCGGCAGAGCTGCCAAGGCGAGCTCGCA 600
DB 639 AGGCCACTTCTTTCACACCTCTCATGAACTCGGCAGAGCTGCCAAGGCGAGCTCGCA 698
QY 601 TCCACGTAAGGATCCACAGAGCTCCATGCTCTCTGGGCGCAGTGAGGCTCTG 660
DB 699 TCCACGTAAGGATCCACAGAGCTCCATGCTCTCTGGGCGCAGTGAGGCTCTG 758
QY 661 GTGACCTCAAGGACATGGAGCTGGCAATCAACCCCTGCAAGCAGCAGCTTGGCCTG 720
DB 759 GTGACCTCAAGGACATGGAGCTGGCAATCAACCCCTGCAAGCAGCAGCTTGGCCTG 818
QY 721 GGACTGGCCAAAGAGCAGCTTCTGTGGAGAGCGCAAGAAATACATGCCAGCCTCCC 780
DB 819 GGACTGGCCAAAGAGCAGCTTCTGTGGAGAGCGCAAGAAATACATGCCAGCCTCCC 878
QY 781 AAGGAAGACCTTTCATCTTCCAGATGAGTAAAGTCTTCTTACGCGATAAGCGCTCT 840
DB 879 AAGGAAGACCTTTCATCTTCCAGATGAGTAAAGTCTTCTTACGCGATAAGCGCTCT 938
QY 841 CTTCTGAGGTCCTCCATAGAAATGAAGAACTGAACCTGGAGACCTTGGTGGTGGCAAAA 900
DB 939 CTTCTGAGGTCCTCCATAGAAATGAAGAACTGAACCTGGAGACCTTGGTGGTGGCAAAA 998
QY 901 AAGATGATGCAAAACACCTGGCCATGAAATATCAACACCTACGTCCTCAGCATACTCAAC 960
DB 999 AAGATGATGCAAAACACCTGGCCATGAAATATCAACACCTACGTCCTCAGCATACTCAAC 1058
QY 961 ATGTATCTGCTTTATTTCAAGATGGAAACAATAGGAGGAAACATCAATTCGAATTGTA 1020
DB 1059 ATGTATCTGCTTTATTTCAAGATGGAAACAATAGGAGGAAACATCAATTCGAATTGTA 1118
QY 1021 GGTCTGATTTCTTAGAAGATGAACGCCAGGCTGGTGTGATTAAGTCAACCGCAGACAC 1080
DB 1119 GGTCTGATTTCTTAGAAGATGAACGCCAGGCTGGTGTGATTAAGTCAACCGCAGACAC 1178
QY 1081 ACCTTAAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCAAT 1140
DB 1179 ACCTTAAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCAAT 1238
QY 1141 GACCAAGCCTTCTTACTGACTGTCTGGATATATGTTCTCTGGAAAGAAATAGCCCTGTGAC 1200
DB 1239 GACCACGCCATCTTACTGACTGTCTGGATATATGTTCTCTGGAAAGAAATAGCCCTGTGAC 1298
QY 1201 ACTTTGGGATTTGCACCCATTAAGTGGAAATGTAGTAAATATCCAGCTGCACGATTAAAT 1260
DB 1299 ACTTTGGGATTTGCACCCATTAAGTGGAAATGTAGTAAATATCCAGCTGCACGATTAAAT 1358
QY 1261 GAAGATACAGCTTTGGAGCTGGCCTTCCACCATGGCCCATGAGTCTGGACACAACTTTGGC 1320
DB 1359 GAAGATACAGCTTTGGAGCTGGCCTTCCACCATGGCCCATGAGTCTGGACACAACTTTGGC 1418
QY 1321 ATGATTTCAATGAGAGAGGGAACATGTGTAAGGTCGAGGGCAACATCATGTCCCT 1380
DB 1419 ATGATTTCAATGAGAGAGGGAACATGTGTAAGGTCGAGGGCAACATCATGTCCCT 1478


```
Db 3639 CCCATTGCAGAGAAAGATGCTCTCTGCAAGACTACTTCCACTGGTCTACCTGGTA 3698
Qy 3601 CCCAGACGGGATCTGCAGCCCAAGTTCTACCGCAAGCAGTGCTCTCAAGACTTGTCTCT 3660
Db 3699 CCCAGACGGGATGTGCAGCCCAAGTTCTACGCAAGCAGTGCTCTCAAGACTTGTCTCT 3758
Qy 3661 AAGTCCCAACTTGTGA 3675
Db 3759 AAGTCCCAACTTGTGA 3773

RESULT 5
AAL43654
ID AAL43654 standard; DNA; 3675 BP.
XX AC AAL43654;
XX DT 05-SEP-2002 (first entry)
XX Human MDTs9 protease coding sequence.
XX Human; gene; ds; MDTs9 protease; TGF-beta inhibitor;
XX transforming growth factor-beta inhibitor; chronic renal failure.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 1..3675
XX FT /*tag= a
XX FT /product= "Human MDTs9 protease"
XX PN WO200251998-A1.
XX PD 04-JUL-2002.
XX PF 21-DEC-2001; 2001WO-JP011251.
XX PR 25-DEC-2000; 2000JP-00393372.
XX PA (YAMA ) YAMANOUCHI PHARM CO LTD.
XX PI Yamaji N, Nishimura K, Abe K, Ogino M;
XX WPI; 2002-508888/54.
XX P-PSDB; AAO15254.
XX Protease MDTs9 and encoded polynucleotide, applicable in diagnosis and
XX screening TGF-beta inhibitors for treatment of chronic renal failure.
XX Example 2; Page 41-48; 60pp; Japanese.
XX The invention comprises the amino acid and coding sequence of the human
XX MDTs9 protease. The MDTs9 DNA and protein sequences of the invention are
XX useful for screening for transforming growth factor (TGF)-beta inhibitors
XX and for the treatment of chronic renal failure. The present DNA sequence
XX encodes the human MDTs9 protease of the invention
XX
SQ Sequence 3675 BP; 873 A; 1007 C; 1020 G; 775 T; 0 U; 0 Other;

Query Match
Best Local Similarity 100.0%; Score 3673.4; DB 6; Length 3675;
Matches 3674; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGAAGCCCGCGCGCGGATGGCGGGCTTTGGCGGCGCTGTGGATGCTGTGGCGCAG 60
Db 1 ATGAAGCCCGCGCGCGGATGGCGGGCTTTGGCGGCGCTGTGGATGCTGTGGCGCAG 60
Qy 61 GTGGCGGAGCAGCACCTTCGTCGCGCATGGGACCGCGCAGCGGCGCCCTGGAGCCCG 120
Db 61 GTGGCGGAGCAGCACCTTCGTCGCGCATGGGACCGCGCAGCGGCGCCCTGGAGCCCG 120
Qy 121 AGCGTCCCGCGTCTCTCCACCCCGGAGCGCGCGGCTGGATGGAAAGCGCGAATAT 180
Db 121 AGCGTCCCGCGTCTCTCCACCCCGGAGCGCGCGGCTGGATGGAAAGCGCGAATAT 180
```

```
Db 121 AGCGTCCCGCGTCTCTCCACCCCGGAGCGCGCGGCTGGATGGAAAGCGCGAATAT 180
Qy 181 GACCTGCTCTCTGCTCTACGAGGTTGACACAGGGGGGATTTAGTGTCCTCAATAAATCATG 240
Db 181 GACCTGCTCTCTGCTCTACGAGGTTGACCAAGGGGGGATTTAGTGTCCTCAATAAATCATG 240
Qy 241 CACCATCAGCGCGGAGAGAGACAGTGGCGGTGTCGAGGTTTGAAGTCTCTTCACTTCGG 300
Db 241 CACCATCAGCGCGGAGAGAGACAGTGGCGGTGTCGAGGTTTGAAGTCTCTTCACTTCGG 300
Qy 301 CTGAAAGGCTCCAGGACGACTTCCACGTGGATCTTGAGGACTTCCAGGACCTTAGTGGCT 360
Db 301 CTGAAAGGCTCCAGGACGACTTCCACGTGGATCTTGAGGACTTCCAGGACCTTAGTGGCT 360
Qy 361 CCTGGCTTTATTTGTGCAGACGTTGGGAAAGACAGCACTAAGTCTGTGCAGACTTTACCG 420
Db 361 CCTGGCTTTATTTGTGCAGACGTTGGGAAAGACAGCACTAAGTCTGTGCAGACTTTACCG 420
Qy 421 CCAGAGGACTTCTGTTTCTATCAAGGCTTTTGGCATCACACAGAAACTCTCTCAGTGGCC 480
Db 421 CCAGAGGACTTCTGTTTCTATCAAGGCTTTTGGCATCACACAGAAACTCTCTCAGTGGCC 480
Qy 481 CTTTCAACTGCAAGGCTTTGTACGSCATGATACGAAACAGAGAGGCGAGTACTTCTTA 540
Db 481 CTTTCAAACCTGCCAAGGCTTTGTACGSCATGATACGAAACAGAGAGGCGAGTACTTCTTA 540
Qy 541 AGGCCACTTCTTTCACACCTCTCATGGAACCTCGGCAGAGCTGCCCAAGGCGAGCTCGCCA 600
Db 541 AGGCCACTTCTTTCACACCTCTCATGGAACCTCGGCAGAGCTGCCCAAGGCGAGCTCGCCA 600
Qy 601 TCCACAGTACTGTACAAGAGATCCACAGAGCCCCATGCTCTCTGGGCGCAGTGGCTCTG 660
Db 601 TCCACAGTACTGTACAAGAGATCCACAGAGCCCCATGCTCTCTGGGCGCAGTGGCTCTG 660
Qy 661 GTGACCTCAAGGACATGGGAGCTGGGCACATCAACCCCTGCACAGCAGGCACTTCGCGCTG 720
Db 661 GTGACCTCAAGGACATGGGAGCTGGGCACATCAACCCCTGCACAGCAGGCACTTCGCGCTG 720
Qy 721 GGACTGCCACAAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCCAGCCTCCC 780
Db 721 GGACTGCCACAAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACATGCCCCAGCCTCCC 780
Qy 781 AAGGAAGACCTCTTTCATCTTCCAGATGAGTATTAAGTCTTCTTACGGCATAAGGCTCT 840
Db 781 AAGGAAGACCTCTTTCATCTTCCAGATGAGTATTAAGTCTTCTTACGGCATAAGGCTCT 840
Qy 841 CTTCTGAGGTCCCATAGAAATGAAGAACTGAACCTGGAGACCTTTGGTGGTGTGACAAA 900
Db 841 CTTCTGAGGTCCCATAGAAATGAAGAACTGAACCTGGAGACCTTTGGTGGTGTGACAAA 900
Qy 901 AAGATGATGCAAAACCATGGCCATGAAATATCACCACTTACGTGCTCAGATACTCAAC 960
Db 901 AAGATGATGCAAAACCATGGCCATGAAATATCACCACTTACGTGCTCAGATACTCAAC 960
Qy 961 ATGTATCTGCTTTATTTCAAAGATGGAAACAATAGGAGGAAACATCAACATTGCAATTGTA 1020
Db 961 ATGTATCTGCTTTATTTCAAAGATGGAAACAATAGGAGGAAACATCAACATTGCAATTGTA 1020
Qy 1021 GGTCTGATTTCTTGAAGATGAACACCCAGGACTGGTGAATAGTCAACACGACAGACAC 1080
Db 1021 GGTCTGATTTCTTGAAGATGAACACCCAGGACTGGTGAATAGTCAACACGACAGACAC 1080
Qy 1081 ACCTTAAGTACTTCTGCCAGTGCAGTCTGGATGATGGGAAAGATGGGACTCGTCAT 1140
Db 1081 ACCTTAAGTACTTCTGCCAGTGCAGTCTGGATGATGGGAAAGATGGGACTCGTCAT 1140
Qy 1141 GACCAACCCCATCTTGACTGGTCTGGATATATGTTCTCTGGAAGATAGGCCCTGTGAC 1200
Db 1141 GACCAACCCCATCTTGACTGGTCTGGATATATGTTCTCTGGAAGATAGGCCCTGTGAC 1200
Qy 1201 ACTTTGGGATTTGCACCCCATAGTGGATGTGTAGTAATATCGCAGCTGCACGATTAAT 1260
Db 1201 ACTTTGGGATTTGCACCCCATAGTGGATGTGTAGTAATATCGCAGCTGCACGATTAAT 1260
```

QY	1261	GAAGATACAGGTCCTTGGAAGGACATGTTGAGTCTGGACACAACTTTGGC	1320
DB	1261	GAAGATACAGGTCCTTGGAAGGACATGTTGAGTCTGGACACAACTTTGGC	1320
QY	1321	ATGATTCATGATGGAAGGACATGTTGAGTCTGGACACAACTTTGGC	1380
DB	1321	ATGATTCATGATGGAAGGACATGTTGAGTCTGGACACAACTTTGGC	1380
QY	1381	ACATTTGGCAGGACGCAATGAGTCTTCTGCTGATCACTTGAAGGACATGTTGGC	1440
DB	1381	ACATTTGGCAGGACGCAATGAGTCTTCTGCTGATCACTTGAAGGACATGTTGGC	1440
QY	1441	AAATTTCTAAGCAACCGCTCAAGTCTTCTGCTGATCACTTGAAGGACATGTTGGC	1500
DB	1441	AAATTTCTAAGCAACCGCTCAAGTCTTCTGCTGATCACTTGAAGGACATGTTGGC	1500
QY	1501	TACAGTATCTGAGAAATGTCAGAGAAATATATGATGCAAAACACAGTGAAGTG	1560
DB	1501	TACAGTATCTGAGAAATGTCAGAGAAATATATGATGCAAAACACAGTGAAGTG	1560
QY	1561	CAGTTGGGAGGAGGACCAAGTCTGATGCTGGAATTTAAAGGACATGTTAAAGCC	1620
DB	1561	CAGTTGGGAGGAGGACCAAGTCTGATGCTGGAATTTAAAGGACATGTTAAAGCC	1620
QY	1621	CTGTGTGATGATGTTGAAGGAAATGTGAGACTAAATTTATGTCAGAGGAGGAGC	1680
DB	1621	CTGTGTGATGATGTTGAAGGAAATGTGAGACTAAATTTATGTCAGAGGAGGAGC	1680
QY	1681	ACAAATTTGTTGGGATGATGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1740
DB	1681	ACAAATTTGTTGGGATGATGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1740
QY	1741	GGCCCCAAGCCCAAGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1800
DB	1741	GGCCCCAAGCCCAAGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1800
QY	1801	ACTGTGGGAGGAGGAGTCTCATAGGAGTGTGCTGCTGCTGCTGCTGCTGCTG	1860
DB	1801	ACTGTGGGAGGAGGAGTCTCATAGGAGTGTGCTGCTGCTGCTGCTGCTGCTG	1860
QY	1861	GGAGGAGGAGTCTGTTGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1920
DB	1861	GGAGGAGGAGTCTGTTGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1920
QY	1921	CCCCGGGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1980
DB	1921	CCCCGGGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1980
QY	1981	AGAGGGGCGGACCTACAAAGTGGAGGAGTCTGCTGCTGCTGCTGCTGCTGCT	2040
DB	1981	AGAGGGGCGGACCTACAAAGTGGAGGAGTCTGCTGCTGCTGCTGCTGCTGCT	2040
QY	2041	CTCTACTGATGTCAGAGGAGTGTGATTTCTTCTTCTTCTTCTTCTTCTTCTT	2100
DB	2041	CTCTACTGATGTCAGAGGAGTGTGATTTCTTCTTCTTCTTCTTCTTCTTCTT	2100
QY	2101	GGGACTCTCATGCTCGAGGAGTGTGATTTCTTCTTCTTCTTCTTCTTCTTCT	2160
DB	2101	GGGACTCTCATGCTCGAGGAGTGTGATTTCTTCTTCTTCTTCTTCTTCTTCT	2160
QY	2161	GGATGTGACAAATGCTTGGATGCTGATGCTGATGCTGATGCTGATGCTGATG	2220
DB	2161	GGATGTGACAAATGCTTGGATGCTGATGCTGATGCTGATGCTGATGCTGATG	2220
QY	2221	AATACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2280
DB	2221	AATACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2280
QY	2281	TATTATCAGATGCTCAGGAGTGTGATTTCTTCTTCTTCTTCTTCTTCTTCTT	2340
DB	2281	TATTATCAGATGCTCAGGAGTGTGATTTCTTCTTCTTCTTCTTCTTCTTCTT	2340

QY	2341	GTCTCTACCTCTCTACATTTCTGTCGCAATGCTGTCAGAAAGGTACTACTGTAATGGGCAC	2400
DB	2341	GTCTCTACCTCTCTACATTTCTGTCGCAATGCTGTCAGAAAGGTACTACTGTAATGGGCAC	2400
QY	2401	TGGACCGTGGACTGGCCCGGCGGTACAAATTTTGGGGCACTACTTTTTCAGACTACAGACGG	2460
DB	2401	TGGACCGTGGACTGGCCCGGCGGTACAAATTTTGGGGCACTACTTTTTCAGACTACAGACGG	2460
QY	2461	TCTATAATGAGCCCGAGAACTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2520
DB	2461	TCTATAATGAGCCCGAGAACTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2520
QY	2521	GAGCTGCTGTTTTCAGGGAAGGAAACCCGGGTGCTGCTGCTGCTGCTGCTGCTGCTG	2580
DB	2521	GAGCTGCTGTTTTCAGGGAAGGAAACCCGGGTGCTGCTGCTGCTGCTGCTGCTGCTG	2580
QY	2581	GGGACCGAAGCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2640
DB	2581	GGGACCGAAGCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2640
QY	2641	TGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2700
DB	2641	TGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2700
QY	2701	AAGTTTCAAGTAAATATATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2760
DB	2701	AAGTTTCAAGTAAATATATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2760
QY	2761	CTTTGCAAAAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2820
DB	2761	CTTTGCAAAAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2820
QY	2821	CGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2880
DB	2821	CGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2880
QY	2881	GACTGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2940
DB	2881	GACTGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2940
QY	2941	TGCMACTCTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3000
DB	2941	TGCMACTCTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3000
QY	3001	ACTGTGGGAAAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	3060
DB	3001	ACTGTGGGAAAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	3060
QY	3061	AGAGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3120
DB	3061	AGAGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3120
QY	3121	GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3180
DB	3121	GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3180
QY	3181	TCCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3240
DB	3181	TCCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3240
QY	3241	AAGTATGTTTTCGGAAGGATTCGAGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTG	3300
DB	3241	AAGTATGTTTTCGGAAGGATTCGAGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTG	3300
QY	3301	CCCAGCTGAGGAGTGGAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3360
DB	3301	CCCAGCTGAGGAGTGGAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3360
QY	3361	GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3420
DB	3361	GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3420
QY	3421	AGCTGTGGGGGAGGCTTCAGACGAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTG	3480

```
Db 3421 AGCTGTGGGAGGCGTTCAGACGAGTCCGTGCGAGTCCCTGCTGGGGCCCGCGGCC 3480
Qy 3481 TCAGGCTGCCCTCTGCAACGAGCCTTCGCGCTCCCTGCGCTGCAACACTCACTTCTGC 3540
Db 3481 TCAGGCTGCCCTCTGCAACGAGCCTTCGCGCTCCCTGCGCTGCAACACTCACTTCTGC 3540
Qy 3541 CCATTGCGAGAGAAAGATGCTTCTGCAAGAGCTTCCCACTGGTGTACCTGGTA 3600
Db 3541 CCATTGCGAGAGAAAGATGCTTCTGCAAGAGCTTCCCACTGGTGTACCTGGTA 3600
Qy 3601 CCCAGCAGGGATGTGCAAGCAGCAAGTCTTACGGCAAGCAGTGTGCAAGACTTCTCT 3660
Db 3601 CCCAGCAGGGATGTGCAAGCAGCAAGTCTTACGGCAAGCAGTGTGCAAGACTTCTCT 3660
Qy 3661 AAGTCCAACTTGTGA 3675
Db 3661 AAGTCCAACTTGTGA 3675

RESULT 6
ACC85689
ID ACC85689 standard; DNA; 3675 BP.
AC ACC85689;
XX
DT 22-APR-2004 (first entry)
XX
DE Human protease MDRS9 related coding sequence.
XX
KW Promoter; human; protease; MDRS9; chronic kidney failure; nephrotropic;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200401047-A1.
XX
PD 31-DEC-2003.
XX
PF 19-JUN-2003; 2003WO-JP007807.
XX
PR 20-JUN-2002; 2002JP-00180543.
XX
PA (YAMA ) YAMANOUCHI PHARM CO LTD.
XX
PI Takeda M, Abe K, Yamaji N;
XX
DR WPI: 2004-071739/07.
DR P-PSDB; ABM79745.
XX
PT New transforming growth factor-beta derived protease MDRS 9 promoter, for
PT screening remedies and/or preventives for chronic kidney failure.
XX
PS Disclosure; Page 30-39; Opp; Japanese.
XX
CC The present invention relates to a polynucleotide, which has a promoter
CC activity. The sequence is derived from the protease MDRS9 gene. The
CC promoter is useful in screening remedies and/or preventives for chronic
CC kidney failure. The present sequence is a polynucleotide shown in the
CC exemplification of the invention
XX
SQ Sequence 3675 BP; 873 A; 1007 C; 1020 G; 775 T; 0 U; 0 Other;

Query Match 100.0%; Score 3673.4; DB 12; Length 3675;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3674; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGAAGCCCGCGCGCGGATGGCGGCTTGGCGGCTGTGTGATGCTGCGCGAG 60
Db 1 ATGAAGCCCGCGCGCGGATGGCGGCTTGGCGGCTGTGTGATGCTGCGCGAG 60
Qy 61 GTGGCCGAGCAGCAGCTTGGTGGCCATGGACCCGCGAGCGGCGCTGGGAGCCG 120
Db 1141 GACCACGCCATCTTACTGACTGTGGATATATGTTCTTCTGGGAAGATGAGCCCTGTGAC 1200
```

```
Db 61 GTGGCCGAGCAGCAGCTTGGTGGCCATGGGACCCGCGAGCGCGCTTGGGAGCCG 120
Qy 121 AGCGTCCCGCGTCTCTCTCCACCCTCCGAGCGCGCGGCTGATGAAAGGCGCAATAT 180
Db 121 AGCGTCCCGCGTCTCTCTCCACCCTCCGAGCGCGCGGCTGATGAAAGGCGCAATAT 180
Qy 181 GACCTGGTCTCTGCGCTACGAGTTGACACAGGGGCGATTACGTGTCCCATGAATCATG 240
Db 181 GACCTGGTCTCTGCGCTACGAGTTGACACAGGGGCGATTACGTGTCCCATGAATCATG 240
Qy 241 CACCATCAGCGCGAGAGAGCAGTGGCGGTCCGAGGTTGAGTCTCTTTCACCTTCGG 300
Db 241 CACCATCAGCGCGAGAGAGCAGTGGCGGTCCGAGGTTGAGTCTCTTTCACCTTCGG 300
Qy 301 CTGAAGGCTCCAGGACGACTTCCACGTTGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360
Db 301 CTGAAGGCTCCAGGACGACTTCCACGTTGATCTGAGGACTTCCAGCAGCCTAGTGGCT 360
Qy 361 CCTGGCTTTATTTGTCAGAGCTTGGGAAAGACAGCACTAAGTCTGTGAGACTTTACCG 420
Db 361 CCTGGCTTTATTTGTCAGAGCTTGGGAAAGACAGCACTAAGTCTGTGAGACTTTACCG 420
Qy 421 CCAGAGGACTTCTGTCTTATCAAGGCTTTCAGGCTTTCGATCACAAGAACTCTCTCAGTGGC 480
Db 421 CCAGAGGACTTCTGTCTTATCAAGGCTTTCGATCACAAGAACTCTCTCAGTGGC 480
Qy 481 CTTTCAACCTGCAAGGCTTGTGAGGATGATAGCAAGAGAGGAGACTTACTTCTCTA 540
Db 481 CTTTCAACCTGCAAGGCTTGTGAGGATGATAGCAAGAGAGGAGACTTACTTCTCTA 540
Qy 541 AGGCACCTCTTTCACACCTCTCATGGAACCTCCGCGAGCTGCCAAGGAGCTGCGCA 600
Db 541 AGGCACCTCTTTCACACCTCTCATGGAACCTCCGCGAGCTGCCAAGGAGCTGCGCA 600
Qy 601 TCCACAGTACTGTACAAGAGATCCACAGAGCCCATGCTCTCTGGGGCCAGTGAGTCTCTG 660
Db 601 TCCACAGTACTGTACAAGAGATCCACAGAGCCCATGCTCTCTGGGGCCAGTGAGTCTCTG 660
Qy 661 GTGACCTCAGGACATGGAGCTGGCACAATCAACCCCTGCAAGAGGAGCTTGGCGCTG 720
Db 661 GTGACCTCAGGAGATGGAGCTGGCACAATCAACCCCTGCAAGAGGAGCTTGGCGCTG 720
Qy 721 GGAAGTCCACAAAGAGCAGCTTCTGTGGAAGAGCGCAAGAAATACATGSCCCAGCTCCC 780
Db 721 GGAAGTCCACAAAGAGCAGCTTCTGTGGAAGAGCGCAAGAAATACATGSCCCAGCTCCC 780
Qy 781 AAGGAAGACCTTCTCATCTTCCAGATGAGTATAGTCTTGTCTTACGCGATAAGCGTCT 840
Db 781 AAGGAAGACCTTCTCATCTTCCAGATGAGTATAGTCTTGTCTTACGCGATAAGCGTCT 840
Qy 841 CTTCTGAGTCCCATAGAAATGAAAGTGAACCTGAGGAGCTTGGTGGTGGTGGTGGTGG 900
Db 841 CTTCTGAGTCCCATAGAAATGAAAGTGAACCTGAGGAGCTTGGTGGTGGTGGTGGTGG 900
Qy 901 AAGATGATGCAAAACCATGGCCATGAAATATACACCTACGTCCTCACGATCTCAAC 960
Db 901 AAGATGATGCAAAACCATGGCCATGAAATATACACCTACGTCCTCACGATCTCAAC 960
Qy 961 ATGATATCTGCTTTATTTCAAGATGGAACAATAGGAGGAAACATCAATGCAATTGTA 1020
Db 961 ATGATATCTGCTTTATTTCAAGATGGAACAATAGGAGGAAACATCAATGCAATTGTA 1020
Qy 1021 GGTCTGATTTCTTAGAAGATGAACAGCAGGAGCTGTGATAGTACACGAGGAGCCAC 1080
Db 1021 GGTCTGATTTCTTAGAAGATGAACAGCAGGAGCTGTGATAGTACACGAGGAGCCAC 1080
Qy 1081 ACCTTAAGTACTTCTGCGAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCAT 1140
Db 1081 ACCTTAAGTACTTCTGCGAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTCGTCAT 1140
Qy 1141 GACCGCCCATCTTACTGACTGTGGATATATGTTCTTCTGGGAAGATGAGCCCTGTGAC 1200
Db 1141 GACCGCCCATCTTACTGACTGTGGATATATGTTCTTCTGGGAAGATGAGCCCTGTGAC 1200
```


Db 3361 GCTGCTGCGGAGCCCTCGAGGGGACGCTGTTGGCTCACCCTGGTCTCAGTGCAGCGCC 3420
 Qy 3421 AGCTGTGGGGAGCGCTTCAGACGAGGTCCTGTCAGTGCCTGGGCGCGCGCGCC 3480
 Db 3421 AGCTGTGGGGAGCGCTTCAGACGAGGTCCTGTCAGTGCCTGGGCGCGCGCGCC 3480
 Qy 3481 TCAGGCTGCTCTCTGCACAGAGCCTTCGGCCCTCCCTGGCTGCAACACTCCTCTCTGC 3540
 Db 3481 TCAGGCTGCTCTCTGCACAGAGCCTTCGGCCCTCCCTGGCTGCAACACTCCTCTCTGC 3540
 Qy 3541 CCATTGTCAGAGAGAGAGATGCTTCTCAAGACTACTTCCACTGGTGTACTCTGGTA 3600
 Db 3541 CCATTGTCAGAGAGAGAGATGCTTCTCAAGACTACTTCCACTGGTGTACTCTGGTA 3600
 Qy 3601 CCCAGCAGCGGATGTGCAGCCACAAAGTCTTACGGCAAGCAGTGTCTGCAAGACTTGTCTCT 3660
 Db 3601 CCCAGCAGCGGATGTGCAGCCACAAAGTCTTACGGCAAGCAGTGTCTGCAAGACTTGTCTCT 3660
 Qy 3661 AAGTCCAACTTGTGA 3675
 Db 3661 AAGTCCAACTTGTGA 3675

RESULT 7
 AAS97174
 ID AAS97174 standard; cDNA; 3675 BP.
 AC AAS97174;
 XX
 DT 26-FEB-2002 (first entry)
 XX
 DE Human metalloprotease partial DNA sequence #3.
 XX

KW Human; protease; PCR primer; cytostatic; immunomodulator; cardiant;
 vasotropic; antimigraine; analgesic; endocrine; neurotropic; tranquiliser;
 hypertensive; hypotensive; neuroleptic; neuroprotective; anabolic;
 anorectic; antiinflammatory; aspartyl protease; cysteine protease;
 metalloprotease; serine protease; cancer; haematopoietic; breast; colon;
 lung; prostate; cervical; brain; ovarian; bladder; kidney; pain;
 immune-related disease; cardiovascular disease; neuronal disease;
 migraine; sexual dysfunction; mood disorder; attention disorder;
 cognition disorder; hypotension; hypertension; psychotic disorder;
 dyskinesia; metabolic disorder; inflammatory disorder; ss.
 XX Homo sapiens.

OS
 XX WO200183782-A2.
 PN
 XX 08-NOV-2001.
 PD
 XX 04-MAY-2001; 2001WO-US014431.
 PF
 XX 04-MAY-2000; 2000US-0201879P.
 PR
 XX (SUGEN-) SUGEN INC.
 PA
 XX Plowman GD, Whyte D, Sudarsanam S, Manning G, Caenepeel S;
 PI Payne V;
 XX
 XX WPI; 2002-041502/05.
 XX P-PSDB; AAU72891.
 DR

PT Novel protease polypeptide useful for screening for substances that may
 be used to treat, e.g., cancers, immune-related diseases, cardiovascular
 PT disease, migraine, pain, psychotic and inflammatory disorders.
 PS Claim 30; Fig 1H; 232pp; English.
 XX
 XX The invention relates to an isolated, enriched, or purified protease
 CC polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to
 CC screen for substances (S) that may modulate its activity. Administering S
 . CC (which modulates protease activity in vitro) may be used to treat a

CC disease or disorder selected from cancers (e.g., of tissues, of blood or
 CC haematopoietic origin, of the breast, colon, lung, prostate, cervical,
 CC brain, ovarian, bladder or kidney), immune-related diseases and
 CC disorders, cardiovascular disease, brain or neuronal-associated diseases
 CC (e.g., central or peripheral nervous system diseases, migraine, pain,
 CC sexual dysfunction, mood disorders, attention disorders, cognitive
 CC disorders, hypotension, hypertension, psychotic disorders, neurological
 CC disorders and dyskinesias), metabolic disorders and inflammatory
 CC disorders. (I) may also be useful as a diagnostic tool for a disease or
 CC disorder such as those above. AAS97159-AAS97195 represent human protease
 CC coding sequences and primers of the invention
 XX

SQ Sequence 3675 BP; 874 A; 1009 C; 1019 G; 773 T; 0 U; 0 Other;

Query Match 99.9%; Score 3671.8; DB 6; Length 3675;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 3673; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAAGCCCGCGCGCGGATGGCGGCGCTTGGCGGCGCTGTGGATGCTGTGGCGGAG 60
 Db 1 ATGAAGCCCGCGCGCGGATGGCGGCGCTTGGCGGCGCTGTGGATGCTGTGGCGGAG 60
 Qy 61 GTGGCCGAGCAGGACCTGCGTGGCCATGGGACCCGCGAGCGCGCTGGGAGCCCG 120
 Db 61 GTGGCCGAGCAGGACCTGCGTGGCCATGGGACCCGCGAGCGCGCTGGGAGCCCG 120
 Qy 121 AGCGTCCCGGCTCTCTCCACCGCGAGCGCGCGCTGGATGGAAGGCGCAATAT 180
 Db 121 AGCGTCCCGGCTCTCTCCACCGCGAGCGCGCGCTGGATGGAAGGCGCAATAT 180
 Qy 181 GACCTGGTCTCTGCTACGAGGTTGACCAAGGCGGCGGATAGTGTCCCATGAATCATG 240
 Db 181 GACCTGGTCTCTGCTACGAGGTTGACCAAGGCGGCGGATAGTGTCCCATGAATCATG 240
 Qy 241 CACCATCAGCGCGGAGAGAGCAGTGGCCGTCTCCGAGGTTGAGTCTCTTCCACCTTCGG 300
 Db 241 CACCATCAGCGCGGAGAGAGCAGTGGCCGTCTCCGAGGTTGAGTCTCTTCCACCTTCGG 300
 Qy 301 CTGAAGGCTCGAGCAGACTTCCACGTGGATCTGAGGACTTCCAGCAGCTAGTGGCT 360
 Db 301 CTGAAGGCTCGAGCAGACTTCCACGTGGATCTGAGGACTTCCAGCAGCTAGTGGCT 360
 Qy 361 CTGCTTTATTGTGACAGCTTGGGAAGACAGGCACTAAGTCTGTGAGACTTTACCG 420
 Db 361 CTGCTTTATTGTGACAGCTTGGGAAGACAGGCACTAAGTCTGTGAGACTTTACCG 420
 Qy 421 CCAGAGGACTTCTGTTCTATCAAGGCTCTTTGGGATCACACAGAACTCTCTAGTGCC 480
 Db 421 CCAGAGGACTTCTGTTCTATCAAGGCTCTTTGGGATCACACAGAACTCTCTAGTGCC 480
 Qy 481 CTTTCACCTGCCAAGGCTTCTCAGGCGATGATACGACAGAGGCGAGATTACTTCTTA 540
 Db 481 CTTTCACCTGCCAAGGCTTCTCAGGCGATGATACGACAGAGGCGAGATTACTTCTTA 540
 Qy 541 AGGCCACTTCTTCCACACTCTCATGGAATCTCGGAGAGCTGCCAAGGAGCTCGCA 600
 Db 541 AGGCCACTTCTTCCACACTCTCATGGAATCTCGGAGAGCTGCCAAGGAGCTCGCA 600
 Qy 601 TCCACGCTACTGTACAGAGATCCACAGAGCCCATGCTCTGGGGCCAGTGAGTCTCTG 660
 Db 601 TCCACGCTACTGTACAGAGATCCACAGAGCCCATGCTCTGGGGCCAGTGAGTCTCTG 660
 Qy 661 GTGACCTCAAGGACTGGGAGCTGGCAATCAACCCCTGACAGCGAGGCTTCGCGCTG 720
 Db 661 GTGACCTCAAGGACTGGGAGCTGGCAATCAACCCCTGACAGCGAGGCTTCGCGCTG 720
 Qy 721 GGACTGCCACAAAAGCAGCACTTCTGTGGAAGAGCAAGAAATACATCCCCAGCCTCCC 780
 Db 721 GGACTGCCACAAAAGCAGCACTTCTGTGGAAGAGCAAGAAATACATCCCCAGCCTCCC 780
 Qy 781 AAGGAAGCCTTCTCATCTTCCAGATGAGTATAGTCTTCTTACGGCATAAAGGCTCT 840
 Db 781 AAGGAAGCCTTCTCATCTTCCAGATGAGTATAGTCTTCTTACGGCATAAAGGCTCT 840

[illegible]

Qy	1921	CCCCGGGACAGTGTGTGACTTCCTCGTGCTGCCTCAGTGTGCCGAGCACAAACAGCACAGATTC	1980
Db	1921	CCCCGGGACAGTGTGTGACTTCCTCGTGCTGCCTCAGTGTGCCGAGCACAAACAGCACAGATTC	1980
Qy	1981	AGAGGGCGGCACACTAACAAGTGGAGCCTTTCACTCAAGTAGAAGATCAGGACTTATGCAA	2040
Db	1981	AGAGGGCGGCACACTAACAAGTGGAGCCTTTCACTCAAGTAGAAGATCAGGACTTATGCAA	2040
Qy	2041	CTCTACTGTATCGCAGAAGGATTGTGATTTCTTCTTTTTCTTTTGTCAAATAAAGTCAAAGAT	2100
Db	2041	CTCTACTGTATCGCAGAAGGATTGTGATTTCTTCTTTTTCTTTTGTCAAATAAAGTCAAAGAT	2100
Qy	2101	GGGACTCATGCTCGGAGGATAGCCGTAATGTTTGTATAGATGGGATATGTGAGAGATT	2160
Db	2101	GGGACTCATGCTCGGAGGATAGCCGTAATGTTTGTATAGATGGGATATGTGAGAGATT	2160
Qy	2161	GGAATGTGACAAATGTCCTTGGATCTGATGCTGTTGAGAGCGTCTGTGGGGTGTCTAACGGG	2220
Db	2161	GGAATGTGACAAATGTCCTTGGATCTGATGCTGTTGAGAGCGTCTGTGGGGTGTGTAAACGGG	2220
Qy	2221	AATAAECTCAGCTGSCACGATTTCAAGGGGTCTCTTACACCAAGCACACCACACCAACCCAG	2280
Db	2221	AATAAECTCAGCTGSCACGATTTCAAGGGGTCTCTTACACCAAGCACACCACACCAACCCAG	2280
Qy	2281	TATTATACATGHTFACCAATTCCTTCTGGAGCCCGGAGTATCCGZATCTATGAAATGAAC	2340
Db	2281	TATTATACATGHTFACCAATTCCTTCTGGAGCCCGGAGTATCCGZATCTATGAAATGAAC	2340
Qy	2341	GTCTCTACTCTTCACATTTCTGTGGCAATGCCCTCAGAGGTACTACCTGAAATGGGCAC	2400
Db	2341	GTCTCTACTCTTCACATTTCTGTGGCAATGCCCTCAGAAAGTACTACCTGAAATGGGCAC	2400
Qy	2401	TGGACCGTGGACTGGCCCGCCGGGTACAAATTTTCGGGCCACTCTTCGACTACAGACGG	2460
Db	2401	TGGACCGTGGACTGGCCCGCCGGGTACAAATTTTCGGGCCACTCTTCGACTACAGACGG	2460
Qy	2461	TCCTTAATAGGCCCGAGAACTTAATGCTTACTGGAACCAACCAACGAGACACTGATTGTG	2520
Db	2461	TCCTTAATAGGCCCGAGAACTTAATGCTTACTGGAACCAACCAACGAGACACTGATTGTG	2520
Qy	2521	GAGCTGTCTGTTTCAAGGAAAGGAACCCGGGTGTTGCCCTGGGAATACTTCATGCTCCCTTG	2580
Db	2521	GAGCTGTCTGTTTCAAGGAAAGGAACCCGGGTGTTGCCCTGGGAATACTTCATGCTCCCTTG	2580
Qy	2581	GGGACCGAGAAGCAGCCCTCTGCCAGCCAGCTACACTTGGGCCCATCGTGCCTCTGAG	2640
Db	2581	GGGACCGAGAAGCAGCCCTCTGCCAGCCAGCTACACTTGGGCCCATCGTGCCTCTGAG	2640
Qy	2641	TGCTCCGTGTCTCTCGGAGGGGACAGATGACCGTGGAGAGGGGTCTACAGAGACCTG	2700
Db	2641	TGCTCCGTGTCTCTCGGAGGGGACAGATGACCGTGGAGAGGGGTCTACAGAGACCTG	2700
Qy	2701	AAGTTTCAAGTAAATATGTCTCTTCTGGCAATCCCAAGACACGACCTGTCAAGGGGTGGTG	2760
Db	2701	AAGTTTCAAGTAAATATGTCTCTTCTGGCAATCCCAAGACACGACCTGTCAAGGGGTGGTG	2760
Qy	2761	CCTTGGAAAAGTATCTGCTCTCTCTCCAGCTGCTCGTGGGAACTGGAGTGCCTCACT	2820
Db	2761	CCTTGGAAAAGTATCTGCTCTCTCTCCAGCTGCTCGTGGGAACTGGAGTGCCTCACT	2820
Qy	2821	CGGACGTGTGGCGGGGTGCCAGAGCCGCCGTGCACTGACACACGCGGGGTGCACTAT	2880
Db	2821	CGGACGTGTGGCGGGGTGCCAGAGCCGCCGTGCACTGCACTGCAACACGCGGGGTGCACTAT	2880
Qy	2881	GACTCGGAGCCAGTCCCGGCACGCTGTGCCCTGTACGCTGTCTCTTCCAGACGAGCGGCC	2940
Db	2881	GACTCGGAGCCAGTCCCGGCACGCTGTGCCCTGTACGCTGTCTCTTCCAGACGAGCGGCC	2940
Qy	2941	TGCAACTCTCAGAGCTGCCACCTGCATGAGCGCCCGGCTGTGGGACAGAGTGTCTCACAC	3000
Db	2941	TGCAACTCTCTCAGAGCTGCCACCTGTATGAGCGCCCGGCTGTGGGACAGAGTGTCTCACAC	3000
Qy	3001	ACCTGTGGGAAAGGGGTGGAGAAAGCGGGCAGTGGCTGTGTAAGAGACCAACCCCTCGGCC	3060

Db 3001 ACCTGTGGGAAGGGTGGAGGAGCGGCGAGTGGCTGTAAAGACACCAACCCCTCGGCC 3060
Qy 3061 AGAGCCGAGTGTGCTGCCGACGCTGTGTGCACTCCGAGCCCAAGCCCGATGCAATGAA 3120
Db 3061 AGAGCCGAGTGTGCTGCCGACGCTGTGTGCACTCCGAGCCCAAGCCCGATGCAATGAA 3120
Qy 3121 GCCTGTCTGCTTCAGCGCTGCCACAAAGCCCAAGAGCTGCAGTGGCTGTGTCGGCTGG 3180
Db 3121 GCCTGTCTGCTTCAGCGCTGCCACAAAGCCCAAGAGCTGCAGTGGCTGTGTCGGCTGG 3180
Qy 3181 TCCCACTGTCTGTGATGTGTAAGAGGAAACACAGAAAAGATTTCTTAAATGTGCTGAA 3240
Db 3181 TCCCACTGTCTGTGATGTGTAAGAGGAAACACAGAAAAGATTTCTTAAATGTGCTGAA 3240
Qy 3241 AAGTATGTTTCTGAAAGTATCGAGAGCTGGCTTCAAGAGTGTCTCAATTTGCGGAAG 3300
Db 3241 AAGTATGTTTCTGAAAGTATCGAGAGCTGGCTTCAAGAGTGTCTCAATTTGCGGAAG 3300
Qy 3301 CCAGGCTGGAGCTGGAACGTGCTCGCGCCCGCTTCCATGCCCAAGGCAACCCCATTT 3360
Db 3301 CCAGGCTGGAGCTGGAACGTGCTCGCGCCCGCTTCCATGCCCAAGGCAACCCCATTT 3360
Qy 3361 GTGTGTGCGGACCTCGAGGGGACAGTGGTTTGCCTCAACCTGTCTCAGTGCACGGCC 3420
Db 3361 GTGTGTGCGGACCTCGAGGGGACAGTGGTTTGCCTCAACCTGTCTCAGTGCACGGCC 3420
Qy 3421 AGCTGTGGGGAGGCGTTGAGAGGCTCGTGCAGTGGCTGTGCTGGGGCCCGCGGGCC 3480
Db 3421 AGCTGTGGGGAGGCGTTGAGAGGCTCGTGCAGTGGCTGTGCTGGGGCCCGCGGGCC 3480
Qy 3481 TCAGGCTGCTCTGTGACCAAGAGCTTCGGGCTTCCCTGCTGCAAACTCACTTCTGCG 3540
Db 3481 TCAGGCTGCTCTGTGACCAAGAGCTTCGGGCTTCCCTGCTGCAAACTCACTTCTGCG 3540
Qy 3541 CCATTGACAGAAAGATGCTTGTGCAAGACTACTTCCACTGTGTGCTGCTGCTGCTGTA 3600
Db 3541 CCATTGACAGAAAGATGCTTGTGCAAGACTACTTCCACTGTGTGCTGCTGCTGCTGTA 3600
Qy 3601 CCCAGCAGCGATGTGACGACCAAGTCTACGCGACGAGTGTGCAAGACTTGTGCTGCT 3660
Db 3601 CCCAGCAGCGATGTGACGACCAAGTCTACGCGACGAGTGTGCAAGACTTGTGCTGCT 3660
Qy 3661 AAGTCCAACTTGTGA 3675
Db 3661 AAGTCCAACTTGTGA 3675

RESULT 8
ACF57047
ID ACF57047 standard; cdNA; 3675 BP.
AC ACF57047;
XX
AC ACF57047;
XX
DT 09-OCT-2003 (first entry)
DE Human ADAMTS16 encoding cdNA SEQ ID NO:1.
KW Human; aggrecanase; enzyme; ADAMTS; ADAMTS16; osteopathic; antiarthritic;
KW antiinflammatory; Gene therapy; osteoarthritis; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3675
FT /tag= a
FT /product= "ADAMTS16"
FT /note= "aggrecanase"
XX
PN WO2003057842-A2.
XX
PD 17-JUL-2003.
XX

PF 30-DEC-2002; 2002WO-US041730.
XX
PR 31-DEC-2001; 2001US-034489SP.
XX
PA (AMHP) WYETH.
XX
PI Agostino MJ, Diblasio E;
XX
DR WPI; 2003-577519/54.
DR P-PSDB; ABR83432.
XX
PT New aggrecanase DNA molecule, useful for treating aggrecanase-associated
PT conditions, including osteoarthritis by inhibiting aggrecanase in a
PT mammal.
XX
PS Claim 1; Fig 1; 54pp; English.
XX
CC The present sequence encodes a human aggrecanase protein designated
CC ADAMTS16. ADAMTS16 has osteopathic, antiinflammatory and antiarthritic
CC activities, and can be used in gene therapy. The aggrecanase protein can
CC be used for identifying inhibitors of aggrecanase activity, which
CC comprises: (a) providing an aggrecanase protein (fragment); (b) combining
CC the aggrecanase with a potential inhibitor; and (c) evaluating whether
CC the potential inhibitor inhibits aggrecanase activity, where the
CC aggrecanase protein is used in a three dimensional structural analysis
CC prior to combining with the potential inhibitor, or is used in a computer
CC aided drug design prior to combining with the potential inhibitor. A
CC composition comprising an antibody that binds to a purified aggrecanase
CC protein can be used for inhibiting aggrecanase activity in a mammal, for
CC treatment of aggrecanase-associated conditions, e.g. osteoarthritis
XX
SQ Sequence 3675 BP; 874 A; 1009 C; 1019 G; 773 T; 0 U; 0 Other;

Query Match 99.9%; Score 3671.8; DB 9; Length 3675;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3673; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGAAGCCCGCGCGCGGATGCGGGCTTGGCGCGCTGTGGATGCTCTCTGCGCGAG 60
Db 1 ATGAAGCCCGCGCGCGGATGCGGGCTTGGCGCGCTGTGGATGCTCTCTGCGCGAG 60
Qy 61 GTGCGCGAGCAGGACCTCTGCGCCATGGGACCCGCGAGCGGCGCTTGGAGCCCG 120
Db 61 GTGCGCGAGCAGGACCTCTGCGCCATGGGACCCGCGAGCGGCGCTTGGAGCCCG 120
Qy 121 AGCGTCCCGCGTCTCTCTCCACCGCGAGCGCGCGGCTGGATGGAAGGCGGAATAT 180
Db 121 AGCGTCCCGCGTCTCTCTCCACCGCGAGCGCGCGGCTGGATGGAAGGCGGAATAT 180
Qy 181 GACCTGGTCTCTGCTACGAGTTGACACAGGCGCGATACGTGTCCCATGAATCATG 240
Db 181 GACCTGGTCTCTGCTACGAGTTGACACAGGCGCGATACGTGTCCCATGAATCATG 240
Qy 241 CACCATCAGCGCGGAGAGAGCAGTGGCGGTGTCGAGGTTGAGTCTCTTCACTTCGG 300
Db 241 CACCATCAGCGCGGAGAGAGCAGTGGCGGTGTCGAGGTTGAGTCTCTTCACTTCGG 300
Qy 301 CTGAAGGCTCCAGCAGGACTTCCAGTGGATCTGAGGACTTCCAGGAGCTAGTGGCT 360
Db 301 CTGAAGGCTCCAGCAGGACTTCCAGTGGATCTGAGGACTTCCAGGAGCTAGTGGCT 360
Qy 361 CCTGGCTTTATTGTGCAGACGTTGGGAAAGACAGCGCACTAAGTCTGTGCAGACTTTACCG 420
Db 361 CCTGGCTTTATTGTGCAGACGTTGGGAAAGACAGCGCACTAAGTCTGTGCAGACTTTACCG 420
Qy 421 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGCATCAACAGAACTCTCTAGTGGCC 480
Db 421 CCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGCATCAACAGAACTCTCTAGTGGCC 480
Qy 481 CTTTCAACCTGCCAAGGCTTGTGAGGATGATACGACAGAGAGGCGAGGATTTACTTCTTA 540
Db 481 CTTTCAACCTGCCAAGGCTTGTGAGGATGATACGACAGAGAGGCGAGGATTTACTTCTTA 540

Db 2701 AAGTTTCAAGTAAATATGTCCTTCTGCAATCCNAGACAGACCTGTGTCAGGGGCTGGTG 2760
Qy 2761 CTTTGAAAGTATCTGCTGTCTCTCCAGCTGGTTCGTGGGAACTGGAGTGCCTGCGAT 2820
Db 2761 CTTTGAAAGTATCTGCTGTCTCTCCAGCTGGTTCGTGGGAACTGGAGTGCCTGCGAT 2820
Qy 2821 CGGACGTGTGGCGGGGTGCCAGAGCCGCCGCTGCTGAGTGCACAGCGGGTGCACAT 2880
Db 2821 CGGACGTGTGGCGGGGTGCCAGAGCCGCCGCTGCTGAGTGCACAGCGGGTGCACAT 2880
Qy 2881 GACTCGGACCCAGTCCCGGCCAGCTGTGCTCCTCAGCTGTCTCCCTCCAGCAGGAGGCC 2940
Db 2881 GACTCGGACCCAGTCCCGGCCAGCTGTGCTCCTCAGCTGTCTCCCTCCAGCAGGAGGCC 2940
Qy 2941 TCGAACTCTCAGAGTGCCTCCACCTGATGAGAGCGCGGCCCTTGGGACAGAGTGTCAAC 3000
Db 2941 TCGAACTCTCAGAGTGCCTCCACCTGATGAGAGCGCGGCCCTTGGGACAGAGTGTCAAC 3000
Qy 3001 ACTGTGGGAAGGGGTGGAGGAAGCGGGCAGTGGCTGTAAAGACACCAACCCCTCGGCC 3060
Db 3001 ACTGTGGGAAGGGGTGGAGGAAGCGGGCAGTGGCTGTAAAGACACCAACCCCTCGGCC 3060
Qy 3061 AGAGCGCAGCTGCTGCTGACATGTGAAAGAGAAACACAGAAAGATTTCTTAAATGTCTGAA 3120
Db 3061 AGAGCGCAGCTGCTGCTGACATGTGAAAGAGAAACACAGAAAGATTTCTTAAATGTCTGAA 3120
Qy 3121 GCCTGTCTGCTTTCAGGCTGCCACAGCCCAAGAGCTGCTGAGTGGTGGTTCGCCCTGG 3180
Db 3121 GCCTGTCTGCTTTCAGGCTGCCACAGCCCAAGAGCTGCTGAGTGGTGGTTCGCCCTGG 3180
Qy 3181 TCCGAGTGTCTGTGACATGTGAAAGAGAAACACAGAAAGATTTCTTAAATGTCTGAA 3240
Db 3181 TCCGAGTGTCTGTGACATGTGAAAGAGAAACACAGAAAGATTTCTTAAATGTCTGAA 3240
Qy 3241 AAGTATGTTTCTGAAAGATPACGAGCTGGCTCAAGAAAGTGTCTCAATTTGCCGAAG 3300
Db 3241 AAGTATGTTTCTGAAAGATPACGAGCTGGCTCAAGAAAGTGTCTCAATTTGCCGAAG 3300
Qy 3301 CCCAGCTGGAGCTGGAGTGCCTGGCGCCCGCTTCCATGCCCGCCAGCCGCCCATTT 3360
Db 3301 CCCAGCTGGAGCTGGAGTGCCTGGCGCCCGCTTCCATGCCCGCCAGCCGCCCATTT 3360
Qy 3361 GCTGTGCGGGACCTCGAGGGGCGAGCTGGTTGGCTTCCCTCAGTGTCTCAGTGCACGGCC 3420
Db 3361 GCTGTGCGGGACCTCGAGGGGCGAGCTGGTTGGCTTCCCTCAGTGTCTCAGTGCACGGCC 3420
Qy 3421 AGCTGTGGGGAGGGTTCAGAGAGGTCCTGTCAGTGGCTGGCTGGGGGCGCGCGGCC 3480
Db 3421 AGCTGTGGGGAGGGTTCAGAGAGGTCCTGTCAGTGGCTGGCTGGGGGCGCGCGGCC 3480
Qy 3481 TCAGGCTGCTCCTGCACAGAGCTTTCGGGCTCCTGGGCTGCAACACTTCTGTC 3540
Db 3481 TCAGGCTGCTCCTGCACAGAGCTTTCGGGCTCCTGGGCTGCAACACTTCTGTC 3540
Qy 3541 CCCATTGCAGAGAAGAAGATCCCTTCTGCAAGACTACTTCCACTGGTGTACCTGGTA 3600
Db 3541 CCCATTGCAGAGAAGAAGATCCCTTCTGCAAGACTACTTCCACTGGTGTACCTGGTA 3600
Qy 3601 CCCAGCAGCGGATGTGACGCCACAAAGTCTACGGCAGAGCTGTGCAAGACTTGTCT 3660
Db 3601 CCCAGCAGCGGATGTGACGCCACAAAGTCTACGGCAGAGCTGTGCAAGACTTGTCT 3660
Qy 3661 AAGTCCAACTTGTGA 3675
Db 3661 AAGTCCAACTTGTGA 3675

RESULT 9
ID ABK49822
XX ABK49822 standard; cDNA; 5610 BP.
AC ABK49822;
XX

DT 15-JUL-2002 (first entry)
XX Human cDNA encoding ADAMTS protein #1.
DE Human; ss; gene; ADAMTS; cytostatic; antidiabetic; antirheumatic; SNP;
KW antiarthritic; antiulcer; vulnery; neovascularisation; angioma;
KW diabetetic omentopathy; chronic rheumatoid arthritis; gene therapy;
KW refractory skin ulcer; gastric ulcer; post-operative healing failure;
KW reopolyisin-type 2N-metalloprotease domain; disintegrin-like domain; TSP1;
KW thrombospondin type 1 domain; sexual cycle; tumour; 5P-syndrome deletion;
KW chromosome 5p15.2-15.3; Cri-du-chat syndrome;
XX single nucleotide polymorphism.
OS Homo sapiens.
PH Key Location/Qualifiers
FT CDS 770..444
FT /*tag= a
FT /product= "ADAMTS"
FT replace(821,T)
FT /*tag= b
FT /standard_name= "single nucleotide polymorphism"
FT replace(1079,T)
FT /*tag= c
FT /standard_name= "single nucleotide polymorphism"
FT replace(1097,G)
FT /*tag= d
FT /standard_name= "single nucleotide polymorphism"
FT replace(1620,A)
FT /*tag= e
FT /standard_name= "single nucleotide polymorphism"
FT replace(2119,C)
FT /*tag= f
FT /standard_name= "single nucleotide polymorphism"
XX WO200231163-A1.
XX 18-APR-2002.
PD 11-OCT-2001; 2001WO-JP008913.
PF 11-OCT-2000; 2000JP-00311309.
PR 02-APR-2001; 2001JP-00102905.
XX (KAZU-) KAZUSA DNA RES INST FOUND.
PA (MITS-) MITSUBISHI PHARMA CORP.
XX Chara O, Nagase T, Nomura N, Yano K, Murakami K, Yasuda S;
PI Kanzaki K;
XX WPI; 2002-372277/40.
DR P-PSDB; AAU79497.
XX Human brain-originated ADAMTS family polypeptide and encoded gene,
PT applicable in diagnosis and screening compounds for drug compositions in
PT treating diseases due to e.g. neovascularization.
XX Claim 4; Page 135-145; 172pp; Japanese.
XX The invention relates to a polypeptide belonging to the ADAMTS family is
CC selected from sequences appearing as AAU79496, AAU79497 and AAU79499, a
CC protein that contains the polypeptide, a protein having not less than 50%
CC homology with the amino acid sequence of the polypeptides or a
CC polypeptide modified from any of the polypeptides but with some amino
CC acids deleted, substituted, added or inserted. Also included are the
CC polynucleotides encoding the polypeptides for their complementary strands
CC or variants), a recombinant vector containing any of the polynucleotides,
CC a transformant which is transformed with the recombinant vector,
CC producing the polypeptide, protein or peptide by culturing the
CC transformant, an antibody that can recognize the polypeptide, protein or
CC peptide and screening compounds to promote or inhibit activity of the
CC polypeptide or protein, or to promote or inhibit expression of the
CC polynucleotide by using the polypeptide, protein, peptide, the

polynucleotide, vector, transformant or/and antibody, particularly in the presence of a test compound for contact before evaluating the activity by measuring signal changes. The polypeptide and encoded gene are applicable in diagnosis and screening compounds for drug compositions in treating diseases due to neovascularisation, diabetic omentopathy, chronic rheumatoid arthritis, angiodysplasia, refractory skin and gastric ulcers and post-operative healing failure, including gene therapy. The gene encoding such polypeptide has conserved repressin-type 2N metalloprotease domain, disintegrin-like domain and TSPI (thrombospondin type 1) domain. Its encoded protein is characterised by high expression in ovaries, changes in expression dose depending on the sexual cycle, a decrease in tumour cell and location of the gene on the 5p-syndrome deletion site on chromosome 5p15.2-15.3 (associated with Cri-du-chat syndrome). The present sequence encodes a ADAMTS protein of the invention

Sequence 5610 BP; 1233 A; 1638 C; 1537 G; 1202 T; 0 U; 0 Other;

Query Match 99.9%; Score 3671.8; DB 6; Length 5610;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3673; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGAAGCCCGCGCGCGCGATGGCGGCTTGGCGCGCTGTGGATGCTGTGGCGCAG 60
770 ATGAAGCCCGCGCGCGCGATGGCGGCTTGGCGCGCTGTGGATGCTGTGGCGCAG 829
61 GTGGCGGAGCGCACCTGCTGCGGCATGGGACCGCGGCGCGCTGGGAGCCCG 120
830 GTGGCGGAGCGCACCTGCTGCGGCATGGGACCGCGGCGCGCTGGGAGCCCG 889
121 AGCGTCCCGCGCTCTCTCCACCGCGGAGCGCGCGGCTGGATGGAAAGCGCAATAT 180
890 AGCGTCCCGCGCTCTCTCCACCGCGGAGCGCGCGGCTGGATGGAAAGCGCAATAT 949
181 GACCTGGTCTCTGCTTACAGGTTGACACAGGGCGGATACGTGTCCCATGAATCATG 240
950 GACCTGGTCTCTGCTTACAGGTTGACACAGGGCGGATACGTGTCCCATGAATCATG 1009
241 CACCATCAGCGCGGAGAGAGAGTGGCGGCTGTGCGAGGTTGAGTCTCTTCACTTGG 300
1010 CACCATCAGCGCGGAGAGAGAGTGGCGGCTGTGCGAGGTTGAGTCTCTTCACTTGG 1069
301 CTGAAAGGCTCCAGGCGACCTTCCAGCTGGATCTGAGGACTTCCAGAGCGCTAGTGGCT 360
1070 CTGAAAGGCGCCAGGCGACCTTCCAGCTGGATCTGAGGACTTCCAGAGCGCTAGTGGCT 1129
361 CTGGCTTTATTTGTGAGAGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
1130 CTGGCTTTATTTGTGAGAGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1189
421 CCAGAGGACTTCTGTTTCTATCAAGGCTTTTGGGATCAACAGAACTCTCTCAAGTGGCC 480
1190 CCAGAGGACTTCTGTTTCTATCAAGGCTTTTGGGATCAACAGAACTCTCTCAAGTGGCC 1249
481 CTTTCAACCTGCGAGGCTTTGAGGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
1250 CTTTCAACCTGCGAGGCTTTGAGGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1309
541 AGGCGACTTCTCTTCAACCTCTCATGGAATCTGCGAGAGCTGCGGAGAGAGAGAGAGAG 600
1310 AGGCGACTTCTCTTCAACCTCTCATGGAATCTGCGAGAGCTGCGGAGAGAGAGAGAGAG 1369
601 TCCCGAGTACTGTAACAAGAGATCCACAGAGCCCCATGCTCTGCGGCGCAGTGAAGTCTTG 660
1370 TCCCGAGTACTGTAACAAGAGATCCACAGAGCCCCATGCTCTGCGGCGCAGTGAAGTCTTG 1429
661 GTGACTCAGAGCATGGAGCTGGGAGATCAACCCCTGCGAGAGAGAGAGAGAGAGAGAG 720
1430 GTGACTCAGAGCATGGAGCTGGGAGATCAACCCCTGCGAGAGAGAGAGAGAGAGAGAG 1489
721 GGACTGCCACAAAGCAGATTTCTGTGGAAGAGCAGCAAGAAATACATGCCCGAGCTCC 780
1490 GGACTGCCACAAAGCAGATTTCTGTGGAAGAGCAGCAAGAAATACATGCCCGAGCTCC 1549

781 AAGGAGAGCTCTTCTATCTTGGCAGATGAGTATAGTCTTGGCTTACGGCATAGCGCTCT 840
1550 AAGGAGAGCTCTTCTATCTTGGCAGATGAGTATAGTCTTGGCTTACGGCATAGCGCTCT 1609
841 CTTCTGAGGTCCCATAGAAATGAAGAACTGAACGTGGAGACCTTGGTGGTGGTGGCAAA 900
1610 CTTCTGAGGTCCCATAGAAATGAAGAACTGAACGTGGAGACCTTGGTGGTGGTGGCAAA 1669
901 AAGATGATGCAAAACATGGCCCATGAAATATCAACCATGCTGCTACGATACCTCAAC 960
1670 AAGATGATGCAAAACATGGCCCATGAAATATCAACCATGCTGCTACGATACCTCAAC 1729
961 ATGGTATCTGCTTTATTCAGAGATGCAAAATAGGAGGAAACATCAACATTCGATTTGA 1020
1730 ATGGTATCTGCTTTATTCAGAGATGCAAAATAGGAGGAAACATCAACATTCGATTTGA 1789
1021 GGTCTGATTTCTTAGAAGATGAACAGCGAGACCTGGTATAGTCAACACGAGACACAC 1080
1790 GGTCTGATTTCTTAGAAGATGAACAGCGAGACCTGGTATAGTCAACACGAGACACAC 1849
1081 ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGATTTGATGGGAAAGATGGGACTCGTCAAT 1140
1850 ACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGATTTGATGGGAAAGATGGGACTCGTCAAT 1909
1141 GACCAAGCGCATCTTACTGACTGCTGATATATCTTCTGGAAGATGAGCCCTGTGAC 1200
1910 GACCAAGCGCATCTTACTGACTGCTGATATATCTTCTGGAAGATGAGCCCTGTGAC 1969
1201 ACTTTGGGATTTGACCCCATAGTGAATGTGTAGTAAATATCGAGCTGCGACGATTAAT 1260
1970 ACTTTGGGATTTGACCCCATAGTGAATGTGTAGTAAATATCGAGCTGCGACGATTAAT 2029
1261 GAAGATACAGTCTTGGGCTTCCCATTTGATGGGAAAGATGGGACTCGTCAACCTTTGGC 1320
2030 GAAGATACAGTCTTGGGCTTCCCATTTGATGGGAAAGATGGGACTCGTCAACCTTTGGC 2089
1321 ATGATTCATGATGAGAGAGGAGAAATGTGTAAAGATCGAGGGCAACATCATGTCCCT 1380
2090 ATGATTCATGATGAGAGAGGAGAAATGTGTAAAGATCGAGGGCAACATCATGTCCCT 2149
1381 ACATTTGGGAGAGCAATGGAGTCTTCTGCTGTCACCTCGAGCGCGCAGTATCTACAC 1440
2150 ACATTTGGGAGAGCAATGGAGTCTTCTGCTGTCACCTCGAGCGCGCAGTATCTACAC 2209
1441 AAATTTCTAAGCACCGCTCAAGCTATCTGCTTGTGCTGATCAGCCAAAGCTGTGAAGGAA 1500
2210 AAATTTCTAAGCACCGCTCAAGCTATCTGCTTGTGCTGATCAGCCAAAGCTGTGAAGGAA 2269
1501 TACAAATATCTGAGAAATTTGCCAGGAGAAATATATGATGCAAAACACACAGTGCAGTGG 1560
2270 TACAAATATCTGAGAAATTTGCCAGGAGAAATATATGATGCAAAACACACAGTGCAGTGG 2329
1561 CAGTTTCGAGAGAGAGCAAGCTGCTGATGCTGGAGCTTTTAAAGGAGATCTGTAAGGCC 1620
2330 CAGTTTCGAGAGAGAGCAAGCTGCTGATGCTGGAGCTTTTAAAGGAGATCTGTAAGGCC 2389
1621 CTGTTGCTCCATCTGATTTGGAGGAAATGTGAGACTAAATTTATGCCAGAGCAGAGAGCC 1680
2390 CTGTTGCTCCATCTGATTTGGAGGAAATGTGAGACTAAATTTATGCCAGAGCAGAGAGCC 2449
1681 ACAATTTGCGGATGATCATGTGTCGCGGAGAGAGAGTGTGTAATATATGATGATGAA 1740
2450 ACAATTTGCGGATGATCATGTGTCGCGGAGAGAGAGTGTGTAATATATGATGATGAA 2509
1741 GGGCCCAAGAGCCACCTAGGCGCATGCTGCGAGTGTCTTCTGCTCCCATGCTCCAGG 1800
2510 GGGCCCAAGAGCCACCTAGGCGCATGCTGCGAGTGTCTTCTGCTCCCATGCTCCAGG 2569
1801 ACCTGCGAGGAGGAGTATCTATAGGAGTGCCTCTGCAACCAAGCCCAAGCCATCGCAT 1860
2570 ACCTGCGAGGAGGAGTATCTATAGGAGTGCCTCTGCAACCAAGCCCAAGCCATCGCAT 2629
1861 GGAGGGAAGTTCTGTGAGGGGCTCCATCGCATCTCTGAAGCTCTGCAACAGTGCAGAAATGT 1920

QY	1261	GAAGATACAGTCTTGGACCTGGCTTACCAATTTGCCATGAGTCTTGACACAACTTTGGC	1320
DB	5312	GAAGATACAGGTCTTGGACCTGGCTTACCAATTTGCCATGAGTCTTGACACAACTTTGGC	5371
QY	1321	ATGATTCATGATGAGAGGAAACATGTGTAAAGAGTCCGAGGGCAACATCATGTCCCTT	1380
DB	5372	ATGATTCATGATGAGAGGAAACATGTGTAAAGAGTCCGAGGGCAACATCATGTCCCTT	5431
QY	1381	ACATTTGGCAGGACCAATGAGTCTTCTCTGTGTCACTGTGACAGCCGAGATCTACAC	1440
DB	5432	ACATTTGGCAGGACCAATGAGTCTTCTCTGTGTCACTGTGACAGCCGAGATCTACAC	5491
QY	1441	AAATTTCTAAGCACCGCTCAAGCTATCTGCTTGTGATCAGCCAAAGCCTGTGAGGAA	1500
DB	5492	AAATTTCTAAGCACCGCTCAAGCTATCTGCTTGTGATCAGCCAAAGCCTGTGAGGAA	5551
QY	1501	TACAAGTATCTGAGAAATTTGCAGGAGAAATATATGATGCAAAACACAGTGCAGTGG	1560
DB	5552	TACAAGTATCTGAGAAATTTGCAGGAGAAATATATGATGCAAAACACAGTGCAGTGG	5611
QY	1561	CAGTTCCGAGAGAAAGCCAGCTCTGCATGCTGGACTTTAAAAAGGACATCTGTAAAGCC	1620
DB	5612	CAGTTCCGAGAGAAAGCCAGCTCTGCATGCTGGACTTTAAAAAGGACATCTGTAAAGCC	5671
QY	1621	CTGTGTGCCATCGTATTGGAAGGAAATGTGAGACTAAATTTATGCCAGCAGCAGAGGC	1680
DB	5672	CTGTGTGCCATCGTATTGGAAGGAAATGTGAGACTAAATTTATGCCAGCAGCAGAGGC	5731
QY	1681	ACAAATTTGTGGGATGACATGTGTGTCGGGAGGACAGTGTGAAATATGATGAA	1740
DB	5732	ACAAATTTGTGGGATGACATGTGTGTCGGGAGGACAGTGTGAAATATGATGAA	5791
QY	1741	GGCCCCAAGCCACCATCGGACCTGGTCCGACCTGGTCTTCTTGGTCCCACTGCTCCAGG	1800
DB	5792	GGCCCCAAGCCACCATCGGACCTGGTCCGACCTGGTCTTCTTGGTCCCACTGCTCCAGG	5851
QY	1801	ACCTCGGAGGGGGAGTATCTCATAGGAGTCCCTCTGCACCAACCCCAAGCCATCGCAT	1860
DB	5852	ACCTCGGAGGGGGAGTATCTCATAGGAGTCCCTCTGCACCAACCCCAAGCCATCGCAT	5911
QY	1861	GGAGGGAGTCTGTGAGGGTCCACTCGCATCTGAAAGCTCTGCAACAGTCAGAAATGT	1920
DB	5912	GGAGGGAGTCTGTGAGGGTCCACTCGCATCTGAAAGCTCTGCAACAGTCAGAAATGT	5971
QY	1921	CCCCGGGACAGTGTGACCTTCCGTGCTCTCAGTGTGCCGACACACAGCAGACGATTC	1980
DB	5972	CCCCGGGACAGTGTGACCTTCCGTGCTCTCAGTGTGCCGACACACAGCAGACGATTC	6031
QY	1981	AGAGGGCGGCACTACAAGTGGAGCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAA	2040
DB	6032	AGAGGGCGGCACTACAAGTGGAGCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAA	6091
QY	2041	CTTACTGTATCGCAGAGGATTTGATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT	2100
DB	6092	CTTACTGTATCGCAGAGGATTTGATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT	6151
QY	2101	GGGACTCCATGCTCGGAGGATAGCCGTAATTTGTATAGATGGATATGTGAGAGATT	2160
DB	6152	GGGACTCCATGCTCGGAGGATAGCCGTAATTTGTATAGATGGATATGTGAGAGATT	6211
QY	2161	GGATGTGACAAATGTCTTGGATCTGATGTGTGAAAGACGTCGTGTGGGTGTAAACGGG	2220
DB	6212	GGATGTGACAAATGTCTTGGATCTGATGTGTGAAAGACGTCGTGTGGGTGTAAACGGG	6271
QY	2221	AATAACTCAGCCTGACGATTCACAGGGGTCTTACACCAAGGACCCACCAACCAACCG	2280
DB	6272	AATAACTCAGCCTGACGATTCACAGGGGTCTTACACCAAGGACCCACCAACCAACCG	6331
QY	2281	TATTAATCATGGTCAACCACTCTTCTGGAGCCGGAGTATCCGATCTATGAAATGAAC	2340
DB	6332	TATTAATCATGGTCAACCACTCTTCTGGAGCCGGAGTATCCGATCTATGAAATGAAC	6391

QY	2341	GTCTCTACCTCTCTACATTTCTGTGGGCAATGCCCTCAGAAAGGTACTACTCTGAATGGGCAC	2400
DB	6392	GTCTCTACCTCTCTACATTTCTGTGGGCAATGCCCTCAGAAAGGTACTACTCTGAATGGGCAC	6451
QY	2401	TGGACCGTGGACTGGCCCGCGCGGTACAAATTTTTCGGGCACCTACTTTTCGACTACAGACGG	2460
DB	6452	TGGACCGTGGACTGGCCCGCGCGGTACAAATTTTTCGGGCACCTACTTTTCGACTACAGACGG	6511
QY	2461	TCCTATAATGAGCCGAGAACTTAATCGCTACTTGGACCAACCAACGAGACACTGATTTGG	2520
DB	6512	TCCTATAATGAGCCGAGAACTTAATCGCTACTTGGACCAACCAACGAGACACTGATTTGG	6571
QY	2521	GAGCTGCTGTTTTCAGGGAAGGAAACCCGGGTGTGGCTGGGAATACTCCATGCTCGCTTG	2580
DB	6572	GAGCTGCTGTTTTCAGGGAAGGAAACCCGGGTGTGGCTGGGAATACTCCATGCTCGCTTG	6631
QY	2581	GGGACCGGAGAGCAGCCCTGCGCCAGCCAGCTACACTTGGGCCATCGTGGCTCTCGAG	2640
DB	6632	GGGACCGGAGAGCAGCCCTGCGCCAGCCAGCTACACTTGGGCCATCGTGGCTCTCGAG	6691
QY	2641	TGCTCCGTGTCTGCGAGGGGGACAGATGACCGTGGAGAGGGCTGTACAGAGACCTG	2700
DB	6692	TGCTCCGTGTCTGCGAGGGGGACAGATGACCGTGGAGAGGGCTGTACAGAGACCTG	6751
QY	2701	AAGTTTCAAGTAAATATGCTCTTCCAAATCCCAAGCACGACCTGTCAACGGGCTGGT	2760
DB	6752	AAGTTTCAAGTAAATATGCTCTTCCAAATCCCAAGCACGACCTGTCAACGGGCTGGT	6811
QY	2761	CCTTGCAGAAATATGCTCTTCCCAAGCTGGTCCGTGGGGAACTGGAGTGCCTGCACT	2820
DB	6812	CCTTGCAGAAATATGCTCTTCCCAAGCTGGTCCGTGGGGAACTGGAGTGCCTGCACT	6871
QY	2821	CGGAGCTGTGGCGGGGTGCCAGAGCGGCCCTGTGACGTGCACACAGCGGGGTGACATAT	2880
DB	6872	CGGAGCTGTGGCGGGGTGCCAGAGCGGCCCTGTGACGTGCACACAGCGGGGTGACATAT	6931
QY	2881	GACTCGGAGCAGTCCCGGCCAGCTGTGCCCTCAGCTGCTGCTCCAGCAGCAGCGCC	2940
DB	6932	GACTCGGAGCAGTCCCGGCCAGCTGTGCCCTCAGCTGCTGCTCCAGCAGCAGCGCC	6991
QY	2941	TGCAACTCTCAGAGCTGCCACCTGTCATGGAGCGCCGGGCCCTGGGAGAGTGTCTCAC	3000
DB	6992	TGCAACTCTCAGAGCTGCCACCTGTCATGGAGCGCCGGGCCCTGGGAGAGTGTCTCAC	7051
QY	3001	ACCTGTGGGAGGGGTGGAGAGCGGGCAGTGGCCCTGTAAAGACACCAACCCCTCGGCC	3060
DB	7052	ACCTGTGGGAGGGGTGGAGAGCGGGCAGTGGCCCTGTAAAGACACCAACCCCTCGGCC	7111
QY	3061	AGAGCGCAGCTGCTGCCCGACGCTGTCTGCACCTCCGAGCCCAAGCCAGGATGCATGAA	3120
DB	7112	AGAGCGCAGCTGCTGCCCGACGCTGTCTGCACCTCCGAGCCCAAGCCAGGATGCATGAA	7171
QY	3121	GCCTGTCTGCTTACGCGCTGCCACAAGCCCAAGAGCTGACGTGGCTGGTGTCCGCTGG	3180
DB	7172	GCCTGTCTGCTTACGCGCTGCCACAAGCCCAAGAGCTGACGTGGCTGGTGTCCGCTGG	7231
QY	3181	TCCCAAGTGTCTGTGACATGTGAAAGAGGAAACACAGAAAGATTTCTTAAATGTGCTGAA	3240
DB	7232	TCCCAAGTGTCTGTGACATGTGAAAGAGGAAACACAGAAAGATTTCTTAAATGTGCTGAA	7291
QY	3241	AAGTATGTTTCTGAAAGTATCGAGAGCTGGCCTCAAAGAGTGTCTCACATTTGCGGAAG	3300
DB	7292	AAGTATGTTTCTGAAAGTATCGAGAGCTGGCCTCAAAGAGTGTCTCACATTTGCGGAAG	7351
QY	3301	CCGAGCTGAGCTGGAAAGTGTGCTGCGCCCGCTTCATGCCCGGACCCCGCCATTTT	3360
DB	7352	CCGAGCTGAGCTGGAAAGTGTGCTGCGCCCGCTTCATGCCCGGACCCCGCCATTTT	7411
QY	3361	GCTGCTCGGGAGCCCTCGAGGGGAGTGTGTTTGGCTCACCTGCTCAGTGTGACGGCC	3420
DB	7412	GCTGCTCGGGAGCCCTCGAGGGGAGTGTGTTTGGCTCACCTGCTCAGTGTGACGGCC	7471
QY	3421	AGCTGTGGGGAGGGCTTTCAGACAGAGTCCGTGAGTGCCTGTGGGGCGCGCGGCC	3480

Db 7472 AGCTGGGGAGGCGCTTCAGACAGGTCGTCAGTGCCTGGCGGGCGCGGCC 7531
Qy 3481 TCAGGCTGCTCTCTGCACAGAGCCTTCGGCTCTCCCTGGCTTGCACCACTCACTTCTGC 3540
Db 7532 TCAGGCTGCTCTCTGCACAGAGCCTTCGGCTCTCCCTGGCTTGCACCACTCACTTCTGC 7591
Qy 3541 CCATTGACAGAGAAAGATGCTTCTGCAAGACTACTTCACTGGTGCTACCTGGTA 3600
Db 7592 CCATTGACAGAGAAAGATGCTTCTGCAAGACTACTTCACTGGTGCTACCTGGTA 7651
Qy 3601 CCCAGCAGGATGTCAGCCACAGTCTACGGCAAGCTGCTCAAGACTTCTCT 3660
Db 7652 CCCAGCAGGATGTCAGCCACAGTCTACGGCAAGCTGCTGCAAGACTTCTCT 7711
Qy 3661 AAGTCCAACTTG 3672
Db 7712 AAGTCCAACTTG 7723
RESULT 11
ABK49826
ID ABK49826 standard; DNA; 8505 BP.
XX
AC ABK49826;
XX
DT 15-JUL-2002 (first entry)
XX
DE Plasmid pFastBac1-MS/HT-PJ01256-2.
XX
KW ds; ADAMTS; cytostatic; antidiabetic; antirheumatic; cyclic;
KW antirheumatic; antiulcer; vulnary; neovascularisation; angioma;
KW diabetic omentopathy; chronic rheumatoid arthritis; gene therapy;
KW refractory skin ulcer; gastric ulcer; post-operative healing failure;
KW repolysin-type ZN-metalloprotease domain; disintegrin-like domain; TSPI;
KW thrombospondin type 1 domain; sexual cycle; tumour; 5P-syndrome deletion;
KW chromosome 5p15.2-15.3; Cri-du-chat syndrome; circular;
KW pFastBac1-MS/HT-PJ01256-2.
XX
OS Homo sapiens.
OS Escherichia coli.
OS unidentified baculovirus.
OS Rhesus macaque polyoma virus.
OS Transposon Tn7.
OS Synthetic.
XX
PN WO200231163-A1.
XX
PD 18-APR-2002.
XX
PF 11-OCT-2001; 2001WO-JP008913.
XX
PR 11-OCT-2000; 2000JP-00311309.
PR 02-APR-2001; 2001JP-00102905.
XX
PA (KAZU-) KAZUSA DNA RES INST FOUND.
PA (MITS-) MITSUBISHI PHARMA CORP.
XX
PI Ohara O, Nagase T, Nomura N, Yano K, Murakami K, Yasuda S;
PI Kanzaki K;
XX
WPI; 2002-372277/40.
XX
PT Human brain-originated ADAMTS family polypeptide and encoded gene,
PT applicable in diagnosis and screening compounds for drug compositions in
PT treating diseases due to e.g. neovascularization.
XX
PS Example 20; Page 159-163; 172pp; Japanese.
XX
CC The invention relates to a polypeptide belonging to the ADAMTS family is
CC selected from sequences appearing as AAU79496, AAU79497 and AAU79499, a
CC protein that contains the polypeptide, a protein having not less than 50%
CC homology with the amino acid sequence of the polypeptides or a

CC polypeptide modified from any of the polypeptides but with some amino
CC acids deleted, substituted, added or inserted. Also included are the
CC polynucleotides encoding the polypeptides (or their complementary strands
CC or variants), a recombinant vector containing any of the polynucleotides,
CC a transformant which is transformed with the recombinant vector,
CC producing the polypeptide, protein or peptide by culturing the
CC transformant, an antibody that can recognize the polypeptide, protein or
CC peptide and screening compounds to promote or inhibit activity of the
CC polypeptide or protein, or to promote or inhibit expression of the
CC polynucleotide by using the polypeptide, protein, peptide, the
CC polynucleotide, vector, transformant or/and antibody, particularly in the
CC presence of a test compound for contact before evaluating the activity by
CC measuring signal changes. The polypeptide and encoded gene are applicable
CC in diagnosis and screening compounds for drug compositions in treating
CC diseases due to neovascularisation, diabetic omentopathy, chronic
CC rheumatoid arthritis, angioma, refractory skin and gastric ulcers and
CC post-operative healing failure, including gene therapy. The gene encoding
CC such polypeptide has conserved repolysin-type ZN- metalloprotease
CC domain, disintegrin-like domain and TSPI (thrombospondin type 1) domain.
CC Its encoded protein is characterised by high expression in ovaries,
CC changes in expression dose depending on the sexual cycle, a decrease in
CC tumour cell and location of the gene on the 5P-syndrome deletion site on
CC chromosome 5p15.2-15.3 (associated with Cri-du-chat syndrome). The
CC present sequence is the expression Plasmid pFastBac1-HT-PJ01256 used to
CC express DNA encoding an ADAMTS protein
XX
SQ Sequence 8505 BP; 2086 A; 2188 C; 2164 G; 2067 T; 0 U; 0 Other;
Query Match 99.8%; Score 3668.8; DB 6; Length 8505;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3670; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGAAGCCCGCGCGCGGATGGCGGCTTGGCGGCTGTGGATGCTGTGGCGCAG 60
Db 4122 ATGAAGCCCGCGCGCGGATGGCGGCTTGGCGGCTGTGGATGCTGTGGCGCAG 4181
Qy 61 GTGGCCGAGCAGGACCTGCGTGCATGGACCCGCGAGCGCGCTGGAGCCCCG 120
Db 4182 GTGGCCGAGCAGGACCTGCGTGCATGGACCCGCGAGCGCGCTGGAGCCCCG 4241
Qy 121 AGCGTCCGCGTCTCTCCACCCGCGAGCGCGCGGCTGGATGGAAAGCGCGAATAT 180
Db 4242 AGCGTCCGCGTCTCTCCACCCGCGAGCGCGCGGCTGGATGGAAAGCGCGAATAT 4301
Qy 181 GACCTGGTCTCTGCTTACAGGTTGACACAGCGGCGATTACGTGTCCCATGAATCATG 240
Db 4302 GACCTGGTCTCTGCTTACAGGTTGACACAGCGGCGATTACGTGTCCCATGAATCATG 4361
Qy 241 CACCATCAGCGCGGAGAGAGCAGTGGCGGTGTCGAGGTTGAGTCTTTCACCTTCG 300
Db 4362 CACCATCAGCGCGGAGAGAGCAGTGGCGGTGTCGAGGTTGAGTCTTTCACCTTCG 4421
Qy 301 CTGAAGGCTCCAGCAGCAGCTTCCACGTGGATCTGAGGACTTCCAGCAGCTAGTGCT 360
Db 4422 CTGAAGGCTCCAGCAGCAGCTTCCACGTGGATCTGAGGACTTCCAGCAGCTAGTGCT 4481
Qy 361 CCGTGGCTTTATTTGTCAGAGCTTGGGAAAGACAGGCATTAAGTCTGTGCAGACTTTACCG 420
Db 4482 CCGTGGCTTTATTTGTCAGAGCTTGGGAAAGACAGGCATTAAGTCTGTGCAGACTTTACCG 4541
Qy 421 CCAGAGGACTTCTGTGTTCTATCAAGGCTCTTTGGCGATCACACAGAACTCTTCAGTGGCC 480
Db 4542 CCAGAGGACTTCTGTGTTCTATCAAGGCTCTTTGGCGATCACACAGAACTCTTCAGTGGCC 4601
Qy 481 CTTTCACTGCGCAAGGCTTGTGAGGATGATACGAAACAGAGAGGCGAGTACTTCTTA 540
Db 4602 CTTTCACTGCGCAAGGCTTGTGAGGATGATACGAAACAGAGAGGCGAGTACTTCTTA 4661
Qy 541 AGGCCACTTCTTTCACCTCTCATGGAACCTCGGCGAGGCTGCCCAAGGCGAGCTGCCA 600
Db 4662 AGGCCACTTCTTTCACCTCTCATGGAACCTCGGCGAGGCTGCCCAAGGCGAGCTGCCA 4721
Qy 601 TCCACGTAAGTGTATACAGAGATCCACAGAGGCCCATGTCTCTGGGGCCAGTGAGGTCTG 660

QY 241 CACCATCAGCGCGGAGAAAGCAGCTGGCCGTGTCGAGGTTGAGTCTCTTCACTTCGG 300
DB 241 CACCATCAGCGCGGAGAAAGCAGCTGGCCGTGTCGAGGTTGAGTCTCTTCACTTCGG 300
QY 301 CTGAAAGGCTCCAGGACGACTTCCA CBTGATCTGAGGACTTCCAGCAGCCTTAGTGCT 360
DB 301 CTGAAAGGCTCCAGGACGACTTCCA CBTGATCTGAGGACTTCCAGCAGCCTTAGTGCT 360
QY 361 CCTGGCTTTATTTGTCAGACGCTGGGAAAGCAGGCACTAAGTCTGTGACACTTTACCG 420
DB 361 CCTGGCTTTATTTGTCAGACGCTGGGAAAGCAGGCACTAAGTCTGTGACACTTTACCG 420
QY 421 CCAGAGGACTTCTGTTTCTATCAAGGCTTTTGGGATCACACAGAAACTCTCTCAGTGCC 480
DB 421 CCAGAGGACTTCTGTTTCTATCAAGGCTTTTGGGATCACACAGAAACTCTCTCAGTGCC 480
QY 481 CTTTCAACCTGCCAAGGCTTGTGAGGCATGATAGAAACAGAGAGGCGAGATTACTTCCTA 540
DB 481 CTTTCAACCTGCCAAGGCTTGTGAGGCATGATAGAAACAGAGAGGCGAGATTACTTCCTA 540
QY 541 AGGCCACTTCTTACACCTCTCATGGAACTCCGCGAGGCTGCCAAGGCACTCGCCA 600
DB 541 AGGCCACTTCTTACACCTCTCATGGAACTCCGCGAGGCTGCCAAGGCACTCGCCA 600
QY 601 TCCACAGTACTGTACAAGAGATCCA CAGAGCCCCATGCTCTCGGGGCGAGTGAGTCCTG 660
DB 601 TCCACAGTACTGTACAAGAGATCCA CAGAGCCCCATGCTCTCGGGGCGAGTGAGTCCTG 660
QY 661 GTGACCTCAAGGACATGGGAGTGGCACA TAAACCCCTGCAAGGAGGACCTTCGCCCTG 720
DB 661 GTGACCTCAAGGACATGGGAGTGGCACA TAAACCCCTGCAAGGAGGACCTTCGCCCTG 720
QY 721 GGACTGCCACAAAAGCAGCTTCTGTGGAAGAGCAAGAAATACATCCCCAGCCTCCC 780
DB 721 GGACTGCCACAAAAGCAGCTTCTGTGGAAGAGCAAGAAATACATCCCCAGCCTCCC 780
QY 781 AAGGAAGACCTCTTCATCTTGCCAGATGAGTAA GTCTTGTTACGGCATAAGCGCTCT 840
DB 781 AAGGAAGACCTCTTCATCTTGCCAGATGAGTAA GTCTTGTTACGGCATAAGCGCTCT 840
QY 841 CTTCTGAGGTCCTCCTAGAAATGAAGAA CTGAACGTGGAGACCTTGGTGGTGAGCAAAA 900
DB 841 CTTCTGAGGTCCTCCTAGAAATGAAGAA CTGAACGTGGAGACCTTGGTGGTGAGCAAAA 900
QY 901 AAGATGATGCAAAACCATGGCCATGAA AATACACCACCTAGCTCCACGATACCTCAAC 960
DB 901 AAGATGATGCAAAACCATGGCCATGAA AATACACCACCTAGCTCCACGATACCTCAAC 960
QY 961 ATGGTATCTGCTTTATTCAAAGATGGA ACAAATPAGAGGAAACATCAACATTTGTA 1020
DB 961 ATGGTATCTGCTTTATTCAAAGATGGA ACAAATPAGAGGAAACATCAACATTTGTA 1020
QY 1021 GGTCTGATTTCTTAGAAGATGAACAGC CAGGACTGGTGATAGTCAACGACGACAC 1080
DB 1021 GGTCTGATTTCTTAGAAGATGAACAGC CAGGACTGGTGATAGTCAACGACGACAC 1080
QY 1081 ACCTTAAGTAGCTCTGCCAGTGGCAG CTGATTTGATGGGAAAGATGGGACTCGTCA 1140
DB 1081 ACCTTAAGTAGCTCTGCCAGTGGCAG CTGATTTGATGGGAAAGATGGGACTCGTCA 1140
QY 1141 GACCACGCCATCTTACTGA CTGGATATATGTTCTCTGGAAGAAATGAGCCCTGTGAC 1200
DB 1141 GACCACGCCATCTTACTGA CTGGATATATGTTCTCTGGAAGAAATGAGCCCTGTGAC 1200
QY 1201 ACTTTGGGATTTGCACCCATAGTGGAA TGTGTAGTAATATCGCAGCTGCACGATTAAT 1260
DB 1201 ACTTTGGGATTTGCACCCATAGTGGAA TGTGTAGTAATATCGCAGCTGCACGATTAAT 1260
QY 1261 GAAGATACAGGCTTTGGACTGGCCCTT CACCATTTGCCATGAGCTGGGACACAACCTTTGCG 1320
DB 1261 GAAGATACAGGCTTTGGACTGGCCCTT CACCATTTGCCATGAGCTGGGACACAACCTTTGCG 1320
QY 1321 ATGATTTCATGATGGAGAAAGGGAAC ATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCT 1380

DB 1321 ATGATTTCATGATGGAGAAAGGGAAC ATGTGTAAAAAGTCCGAGGGCAACATCATGTCCCT 1380
QY 1381 ACATTGGCAGCGCAATGAGTCTTCTCTG TGTACCCCTGCAGCCGCACTATCTACAC 1440
DB 1381 ACATTGGCAGCGCAATGAGTCTTCTCTG TGTACCCCTGCAGCCGCACTATCTACAC 1440
QY 1441 AAATTTCTAAGCACCGCTCAAGCTTATCTG CTTCTGATCAGCCAAAGCCTGTGAAGAA 1500
DB 1441 AAATTTCTAAGCACCGCTCAAGCTTATCTG CTTCTGATCAGCCAAAGCCTGTGAAGAA 1500
QY 1501 TACAAGTATCTGAGAAATTTGCCAGGAGA ATTATATGATGCAAAACACACAGTGAAGTGG 1560
DB 1501 TACAAGTATCTGAGAAATTTGCCAGGAGA ATTATATGATGCAAAACACACAGTGAAGTGG 1560
QY 1561 CAGTTCCGAGAGAAAGCCAAAGCTCTGCATG CTGACCTTTAAAAAGGACATCTGTAAGCC 1620
DB 1561 CAGTTCCGAGAGAAAGCCAAAGCTCTGCATG CTGACCTTTAAAAAGGACATCTGTAAGCC 1620
QY 1621 CTGTGTGCCATCGTATTTGGAAGAAATGTG AGACTAAATTTATGCCAGCAGCAGAGGC 1680
DB 1621 CTGTGTGCCATCGTATTTGGAAGAAATGTG AGACTAAATTTATGCCAGCAGCAGAGGC 1680
QY 1681 ACAATTTGTGGGCATGACATGTGTGCCGGG AGGACAGTGTGTGAATATGTTGATGAA 1740
DB 1681 ACAATTTGTGGGCATGACATGTGTGCCGGG AGGACAGTGTGTGAATATGTTGATGAA 1740
QY 1741 GGCCCCAAGCCCCACCTGCTCGGACCTGCTT CTTGTTGCCCATGTCTCCAGG 1800
DB 1741 GGCCCCAAGCCCCACCTGCTCGGACCTGCTT CTTGTTGCCCATGTCTCCAGG 1800
QY 1801 ACCTGCGAGGGGAGTATCTCATAGAGTGC CTCTGCAACACCCCAAGCCATCGCAT 1860
DB 1801 ACCTGCGAGGGGAGTATCTCATAGAGTGC CTCTGCAACACCCCAAGCCATCGCAT 1860
QY 1861 GGAGGGAAGTTCTGTGAGGGCTCCACTGCGC ACTCTGAAAGTCTGCAACAGTCAGAAATGT 1920
DB 1861 GGAGGGAAGTTCTGTGAGGGCTCCACTGCGC ACTCTGAAAGTCTGCAACAGTCAGAAATGT 1920
QY 1921 CCCCAGGACAGTGTGCTCCGTGCTCAGTGTG CCGAGCACAACAGCAGAGATTC 1980
DB 1921 CCCCAGGACAGTGTGCTCCGTGCTCAGTGTG CCGAGCACAACAGCAGAGATTC 1980
QY 1981 AGAGGGGGCACTACAAAGTGGAAAGCCTTAC ACTCAAGTAGAAGATCAGGACTTATGCAAA 2040
DB 1981 AGAGGGGGCACTACAAAGTGGAAAGCCTTAC ACTCAAGTAGAAGATCAGGACTTATGCAAA 2040
QY 2041 CTCTACTGTATCGCAGAAAGGATTTGATTTCT TTTTGTCAAATAAAGTCAAAGAT 2100
DB 2041 CTCTACTGTATCGCAGAAAGGATTTGATTTCT TTTTGTCAAATAAAGTCAAAGAT 2100
QY 2101 GGGACTCCATGCTCGGAGGATAGCCGTAATGT TTTGTATAGATGGGATATGTGAGAGAT 2160
DB 2101 GGGACTCCATGCTCGGAGGATAGCCGTAATGT TTTGTATAGATGGGATATGTGAGAGAT 2160
QY 2161 GGAATGCAAAATGCTCTTGGATCTGATGCTGT TGAAGACGCTGTGGGGTGTGAACCGG 2220
DB 2161 GGAATGCAAAATGCTCTTGGATCTGATGCTGT TGAAGACGCTGTGGGGTGTGAACCGG 2220
QY 2221 AATAACTCAGCTGCACGATTCACAGGGTCTCT ACACCAAGCACCACCAACCCAG 2280
DB 2221 AATAACTCAGCTGCACGATTCACAGGGTCTCT ACACCAAGCACCACCAACCCAG 2280
QY 2281 TATTATACATGCTCACCATTCTTCTGGAGCCG GAGTATCCGATCTATGAATGAAC 2340
DB 2281 TATTATACATGCTCACCATTCTTCTGGAGCCG GAGTATCCGATCTATGAATGAAC 2340
QY 2341 GTCTCTACCTCCTACATTTCTGTGCGCAATGC CTTCAGAAAGTACTACTGAATGGGAC 2400
DB 2341 GTCTCTACCTCCTACATTTCTGTGCGCAATGC CTTCAGAAAGTACTACTGAATGGGAC 2400
QY 2401 TGGACCGTGGACTGGCCCGCGGTTACAAATTT TTTGGGCACCTACTTTTCGACTACAGACGG 2460

Db	721	CTGGCACATCAACCCCTGTCACAGCAGCGACCTTTCGCCCTGGGACTGCCACAAAGACGAT	780
Qy	742	TTCTCTGGAGACGCAAGNAATACATGCCCGAGCCTCCCAAGGAAGACTCTTCACTCTTG	801
Db	781	TTCTGTGGAGACGCAAGNAATACATATGCCCCAGCCTCCCAAGGAAGACTCTTCACTCTTG	840
Qy	802	CCAGATGAGTAAAGTCTTTCCTAGCGGCAATAAGCCCTCTCTTCTGAGGTCCCATAGAAAT	861
Db	841	CCAGATGAGTAAAGTCTTTCCTAGCGGCAATAAGCCCTCTCTTCTGAGGTCCCATAGAAAT	900
Qy	862	GAAGAACTGAACGTGGAGACCTTGTGTGTGTGTCGACAAAGATGATCAAAACCATGGC	921
Db	901	GAAGAACTGAACGTGGAGACCTTGTGTGTGTGTCGACAAAGATGATCAAAACCATGGC	960
Qy	922	CATGAAATATCACACCTACGTGCTCAGGATCTCAACATGGTATCTGCTTTATCAAA	981
Db	961	CATGAAATATCACACCTACGTGCTCAGGATCTCAACATGGTATCTGCTTTATCAAA	1020
Qy	982	GATGAACAATAGGAGGAAACATCAACATTTGTAAGTCTGATTTCTCTAGAAAT	1041
Db	1021	GATGAACAATAGGAGGAAACATCAACATTTGTAAGTCTGATTTCTCTAGAAAT	1080
Qy	1042	GAACAGCAGGACTGGTGAATAGTCAACAGCAGACCAACCTTAAAGTAGCTTCTGCCAG	1101
Db	1081	GAACAGCAGGACTGGTGAATAGTCAACAGCAGACCAACCTTAAAGTAGCTTCTGCCAG	1140
Qy	1102	TGGCAGTCTGGATTGATGGGGAAGATGGGACTCGTCATGACCGCCATCTTACTGACT	1161
Db	1141	TGGCAGTCTGGATTGATGGGGAAGATGGGACTCGTCATGACCGCCATCTTACTGACT	1200
Qy	1162	GGTCTGGATATATGTTCTCTGGAAGAAATGAGCCCTGTGACACTTTGGGATTTGCAACCAT	1221
Db	1201	GGTCTGGATATATGTTCTCTGGAAGAAATGAGCCCTGTGACACTTTGGGATTTGCAACCAT	1260
Qy	1222	AGTGAATGTGTAGTAATAATCGCAGCTGCAGATTAATGAAGATACAGGCTTGGACTG	1281
Db	1261	AGTGAATGTGTAGTAATAATCGCAGCTGCAGATTAATGAAGATACAGGCTTGGACTG	1320
Qy	1282	GCCTTACCATTTGCCATGAGTCTGCACACAACCTTTGGCATGATTCATGATGGAGAGGG	1341
Db	1321	GCCTTACCATTTGCCATGAGTCTGCACACAACCTTTGGCATGATTCATGATGGAGAGGG	1380
Qy	1342	AACATGTGTAAAGTCCGAGGGCAACATCATGTCCCTCATATTTGGCAGGACGCAATGGA	1401
Db	1381	AACATGTGTAAAGTCCGAGGGCAACATCATGTCCCTCATATTTGGCAGGACGCAATGGA	1440
Qy	1402	GTCTTCTCGTGTGCTGACCTCGAGCCGCGCAGTATCTACACAAATTTCTAAGCACCGCTCAA	1461
Db	1441	GTCTTCTCGTGTGCTGACCTCGAGCCGCGCAGTATCTACACAAATTTCTAAGCACCGCTCAA	1500
Qy	1462	GCTATCTGCTTGTGATCAGCACAAGCTGTGGAAGGAATACAAGTATCCTGAGAAATTG	1521
Db	1501	GCTATCTGCTTGTGATCAGCACAAGCTGTGGAAGGAATACAAGTATCCTGAGAAATTG	1560
Qy	1522	CCAGGAAATATATGATGACAAACACACAGTGCAGTGGCAGTTTGGGAGAGAAAGCCAAAG	1581
Db	1561	CCAGGAAATATATGATGACAAACACACAGTGCAGTGGCAGTTTGGGAGAGAAAGCCAAAG	1620
Qy	1582	CTCTGCATGCTGGACTTTAAAGAGACATCTGTAAAGCCCTGTGGTGCCATCGTATTGGA	1641
Db	1621	CTCTGCATGCTGGACTTTAAAGAGACATCTGTAAAGCCCTGTGGTGCCATCGTATTGGA	1680
Qy	1642	AGGAAATGTGAGACTTAAATTTATGCGCAGCAGAGGACACAATTTTGGGGCATGACATG	1701
Db	1681	AGGAAATGTGAGACTTAAATTTATGCGCAGCAGCAGAGGACACAATTTTGGGGCATGACATG	1740
Qy	1702	TGGTCCCGGGAGGACAGTGTGTGAAATATAGTGTATGAAGGCCCAACCCACCATGGC	1761
Db	1741	TGGTCCCGGGAGGACAGTGTGTGAAATATAGTGTATGAAGGCCCAACCCACCATGGC	1800
Qy	1762	CACCTGGTCCGACTGCTTCTTGTGTCCTCCAGGACCTCGGAGGGGGAGTATCT	1821
Db	1801	CACCTGGTCCGACTGCTTCTTGTGTCCTCCAGGACCTCGGAGGGGGAGTATCT	1860
Qy	1822	CATAGGAGTGGCTCTGCACCAACCCCAAGCCATCGCATGGAGGAGTTCTGTGAGGGC	1881
Db	1861	CATAGGAGTGGCTCTGCACCAACCCCAAGCCATCGCATGGAGGAGTTCTGTGAGGGC	1920
Qy	1882	TCCACTCGCACTCTGAAGCTCTGCAACAGTCAAGAAATGTCCTCCGGGACAGTGTGACTTC	1941
Db	1921	TCCACTCGCACTCTGAAGCTCTGCAACAGTCAAGAAATGTCCTCCGGGACAGTGTGACTTC	1980
Qy	1942	CGTGTGCTCAGTGTGGGAGCAACACAGCAGACGATTCAGAGGGCGGCACTACAAGTGG	2001
Db	1981	CGTGTGCTCAGTGTGGGAGCAACACAGCAGACGATTCAGAGGGCGGCACTACAAGTGG	2040
Qy	2002	AAGCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAACTCTACTGTATCGCAAGAGA	2061
Db	2041	AAGCCTTACACTCAAGTAGAAGATCAGGACTTATGCAAACTCTACTGTATCGCAAGAGA	2100
Qy	2062	TTTGATTTCTTCTTTTCTTTTGTCAAATAAAGTCAAAGTGGGACTCCATGTCTCGAGGAT	2121
Db	2101	TTTGATTTCTTCTTTTCTTTTGTCAAATAAAGTCAAAGTGGGACTCCATGTCTCGAGGAT	2160
Qy	2122	AGCCGTAAATGTTGTATAGATGGGATATGTGAGAGAGTTGGATGTGACAAATGCTCTTGA	2181
Db	2161	AGCCGTAAATGTTGTATAGATGGGATATGTGAGAGAGTTGGATGTGACAAATGCTCTTGA	2220
Qy	2182	TCTCATGCTGTGTAAGACGCTCTGTGGGGTGTAAACGGGAATAACTCAGCCTGCACGAT	2241
Db	2221	TCTCATGCTGTGTAAGACGCTCTGTGGGGTGTAAACGGGAATAACTCAGCCTGCACGAT	2280
Qy	2242	CACAGGGGTCTCTACACCAAGCACCCACACCAACCAAGTATTATCATGGTCAACCAT	2301
Db	2281	CACAGGGGTCTCTACACCAAGCACCCACACCAACCAAGTATTATCATGGTCAACCAT	2340
Qy	2302	CTTCTGGAGCCCGAGTATCCGCATCTATGAAATGAACGTCCTACCTCTCATATTTCT	2361
Db	2341	CTTCTGGAGCCCGAGTATCCGCATCTATGAAATGAACGTCCTACCTCTCATATTTCT	2400
Qy	2362	GTGGCAATGCTCTCAGAAGGTACTACTGTAATGGGCACTGCAGCCGTGGACTGGCCCGGC	2421
Db	2401	GTGGCAATGCTCTCAGAAGGTACTACTGTAATGGGCACTGCAGCCGTGGACTGGCCCGGC	2460
Qy	2422	CGGTACAAAATTTTCGGGCACTACTTTTCGACTACAGAGCGTCTATAATAGCCCCGAGAAC	2481
Db	2461	CGGTACAAAATTTTCGGGCACTACTTTTCGACTACAGAGCGTCTATAATAGCCCCGAGAAC	2520
Qy	2482	TTAATCCGTACTGACCAACCAACGACACTGATTTGTGGAGCTGCTGTTTCAGGGAGG	2541
Db	2521	TTAATCCGTACTGACCAACCAACGACACTGATTTGTGGAGCTGCTGTTTCAGGGAGG	2580
Qy	2542	AACCCGGGTGTTGGCTGGGAATACTCCATGCTCGCTTGGGGAACCGAGAGCAGCCCCCT	2601
Db	2581	AACCCGGGTGTTGGCTGGGAATACTCCATGCTCGCTTGGGGAACCGAGAGCAGCCCCCT	2640
Qy	2602	GCCAGCCCCAGCTACACTTGGGCACTCGTGCGCTCTGAGTGTCTCGGTCTCGGTGGAGGG	2661
Db	2641	GCCAGCCCCAGCTACACTTGGGCACTCGTGCGCTCTGAGTGTCTCGGTGGAGGG	2700
Qy	2662	GGACAGATGACCGTGAGAGAGGGCTCTACAGACCTGAGTTTCAAGTAAATATGTGCC	2721
Db	2701	GGACAGATGACCGTGAGAGAGGGCTCTACAGACCTGAGTTTCAAGTAAATATGTGCC	2760
Qy	2722	TTCTGCAATCCCAAGACACGACCTGTCAACGGGGCTGGTGCCTTGCAAAGTATCTGCCTGT	2781
Db	2761	TTCTGCAATCCCAAGACACGACCTGTCAACGGGGCTGGTGCCTTGCAAAGTATCTGCCTGT	2820
Qy	2782	CCTCCAGCTGCTCGGTGGGAACTGAGTGTCTGCAAGTGTGGAGGGGGTGGCC	2841
Db	2821	CCTCCAGCTGCTCGGTGGGAACTGAGTGTCTGCAAGTGTGGAGGGGGTGGCC	2880
Qy	2842	CAGAGCCGCCCGGTGACAGTGCACACGGCGGTGCACTATGACTCGGAGCGGCTCCCGCC	2901
Db	2881	CAGAGCCGCCCGGTGACAGTGCACACGGCGGTGCACTATGACTCGGAGCGGCTCCCGCC	2940

Db		2281	AGCCAGTCCCGGCGCAGCCTTGCCCTCAGCCTGCTCCCTCCAGCAGGCGGCGCTGCAACT	2340
Qy		2948	CTCAGAGCTGCCACCTCGATGAGCGCGGGCCCTGGGCGAGTGTCTCACACACCTGTG	3007
Db		2341	CTCAGAGCTGCCACCTCGATGAGCGCGGGCCCTGGGCGAGTGTCTCACACACCTGTG	2400
Qy		3008	GGAAGGGGTGGAGGAAGCGGGCAGTGGCTGTAAAGAGCAACAAACCCCTCGGCGAGAGCGC	3067
Db		2401	GGAAGGGGTGGAGGAAGCGGGCAGTGGCTGTAAAGAGCAACAAACCCCTCGGCGAGAGCGC	2460
Qy		3068	AGCTGCTGCCGCGAGCGCTGCTGCACCTCCGAGCCCAAGCCCAAGATGCATGAGCCTGTC	3127
Db		2461	AGCTGCTGCCGCGAGCGCTGCTGCACCTCCGAGCCCAAGCCCAAGATGCATGAGCCTGTC	2520
Qy		3128	TGCTTCAGCGCTGCCCAAGACCCCAAGAGCTGCAGTGGTGTCCGCTGCTGCCAGT	3187
Db		2521	TGCTTCAGCGCTGCCCAAGACCCCAAGAGCTGCAGTGGTGTCCGCTGCTGCCAGT	2580
Qy		3188	GCTCTGTGACATGTGAAGAGGAACAACAGAAAGATTTCTTAAATGTCTGAAAGATATG	3247
Db		2581	GCTCTGTGACATGTGAAGAGGAACAACAGAAAGATTTCTTAAATGTCTGAAAGATATG	2640
Qy		3248	TTTCTGGAAGTATCGAGAGCTGGCCTCAAGAGTGTCTCAATTTGCCGAGCCAGCC	3307
Db		2641	TTTCTGGAAGTATCGAGAGCTGGCCTCAAGAGTGTCTCAATTTGCCGAGCCAGCC	2700
Qy		3308	TGAGAGCTGGAACGTGCTCGCCCGCTTCCATGCCCCAGGCACCCGCCATTTGCTGTG	3367
Db		2701	TGAGAGCTGGAACGTGCTCGCCCGCTTCCATGCCCCAGGCACCCGCCATTTGCTGTG	2760
Qy		3368	CGGGACCTTCGAGGGGCGAGTGGTTTGGCTCACCTGGTCTCAGTGCACGGCCAGCTGTG	3427
Db		2761	CGGGACCTTCGAGGGGCGAGTGGTTTGGCTCACCTGGTCTCAGTGCACGGCCAGCTGTG	2820
Qy		3428	GGGGAGGCGTTTCAGACGAGGTCGTGCAGTGCCTGCTGGGGCCGGCCGCGCTCAGGCT	3487
Db		2821	GGGGAGGCGTTTCAGACGAGGTCGTGCAGTGCCTGCTGGGGCCGGCCGCGCTCAGGCT	2880
Qy		3488	GCCTCTGCACCAAGAGCCTTGGCCCTCCCTGGCCTGCAACACTCACTTCTGCCCATTTG	3547
Db		2881	GCCTCTGCACCAAGAGCCTTGGCCCTCCCTGGCCTGCAACACTCACTTCTGCCCATTTG	2940
Qy		3548	CAGAGAAGAAAGATGCTTTCTGCAAGAGACTTCCACTGGTGCTACCTGGTACCCCGAGC	3607
Db		2941	CAGAGAAGAAAGATGCTTTCTGCAAGAGACTTCCACTGGTGCTACCTGGTACCCCGAGC	3000
Qy		3608	ACGGATGTGCGGCCCAAGTTCTACGGCAAGCAGTGTGCAAGATTGCTCTAAGTCCA	3667
Db		3001	ACGGATGTGCGGCCCAAGTTCTACGGCAAGCAGTGTGCAAGATTGCTCTAAGTCCA	3060
Qy		3668	ACTTGTGA	3675
Db		3061	ACTTGTGA	3068

Search completed: June 13, 2005, 21:26:34
Job time : 1757 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2005, 16:34:24 ; Search time 10401 Seconds
(without alignments)
13449.302 Million cell updates/sec

Title: US-10-804-457-3
Perfect score: 3675
Sequence: 1 atgaagcccccgcgcgcggg.....gcttaagtccaactgtga 3675

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2322.8	63.2	5104	3 AK031314	Mus muscu
2	1120.8	30.5	4335	3 AK053851	Mus muscu
3	808.4	22.0	1092	4 BM080410	AGENCOURT
4	795.8	21.7	4593	3 AK036137	Mus muscu
5	724.4	19.7	1122	5 BM906555	AGENCOURT
6	623.4	17.0	625	4 BM850160	K-EST01130
7	621.4	16.9	623	4 BM845044	K-EST0123
8	535.4	14.6	537	4 BM845406	K-EST0123
9	523.8	14.3	728	6 CD355731	UT-M-FY0
10	511.2	13.9	523	6 CD366691	56008692H
11	509.6	13.9	523	6 CD366687	56008668H
12	508	13.8	523	6 CD366689	56008676H
13	507	13.6	522	6 CD366693	56008776H
14	498.6	13.5	523	6 CD366688	56008668J
15	497.8	13.5	630	7 CF170272	B0825B02
16	497.8	13.5	630	7 CF170316	B0825G07
17	474.2	12.9	735	5 BM944992	UT-M-EH0P
18	466.4	12.7	915	4 BI686750	603312988
19	460	12.5	521	6 CD366694	56008776J
20	458	12.5	527	6 CD366690	56008676J
21	426.8	11.6	430	4 BM844919	K-EST0123
22	402	10.9	3399	3 AK035797	Mus muscu
23	401.2	10.9	892	5 BX458327	EX458327
24	396.4	10.8	698	5 BQ444190	UT-M-EX0-

25	394.6	10.7	845	7 CR425675	CR425675
26	372.6	10.1	788	6 CD636667	CD636667
27	329.4	9.0	332	2 BF933693	IL5-NT022
28	328.2	8.9	718	7 CD433858	CD433858
29	324.4	8.8	816	6 CD636670	56100678J
30	320.6	8.7	763	4 BI690351	BI690351
31	310	8.4	436	6 CD636692	56008692J
32	308.4	8.4	732	6 CA512127	CA512127
33	302.4	8.2	842	6 CD636666	56100586J
34	298.8	8.1	693	7 CN459517	UI-M-HN0-
35	298.6	8.1	774	6 CD636672	CD636672
36	294.4	8.0	462	7 CF534948	CF534948
37	280.4	7.6	3231	3 AK048612	Mus muscu
38	269.4	7.3	3249	3 AK048600	Mus muscu
39	262.8	7.2	3525	3 AK048211	Mus muscu
40	260.2	7.1	426	6 CD636652	CD636652
41	260.2	7.1	426	6 CD636664	55100624J
42	259.2	7.1	425	6 CD636662	55100672J
43	258.2	7.0	424	6 CD636658	CD636658
44	258.2	7.0	424	6 CD636660	55100656J
45	256.6	7.0	424	6 CD636650	CD636650

ALIGNMENTS

RESULT 1	AK031314	5104 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK031314	Mus musculus 13 days embryo male testis cDNA, RIKEN full-length			
DEFINITION	AK031314	enriched library, clone:6030405121 product:hypothetical Reprolysin family propeptide containing protein, full insert sequence.			
ACCESSION	AK031314.1	GI:26082263			
VERSION	AK031314.1	HTC; CAP trapper.			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	1	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)				
PUBMED	99279253				
REFERENCE	2	10349636			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kishimoto, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	20530913				
PUBMED	11076861				
REFERENCE	4				
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409, 685-690 (2001)				
REFERENCE	5				
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research				

1376 CCCCTACATTGGCAGACGCAATGGAGTCTTCTCTGGTCAACCTGCGAGCCGCCAGTATC 1435
1593 CCCCACACTGGCAGACGCAATGGTGTCTTCTCTGGTCTTCTGCGAGCCGTCAGTACC 1652
1436 TACACAAATTTCTAAGCACCGGTCAAGCTATCTGCTTGTCTGATCAGCCCAAGACCTGTGA 1495
1653 TGCACAAAGTTCTTAAGCACCGCCCAAGCGATATGTCTTCTGATCAGCCCAAGACCTGTGA 1712
1496 AGGAATACAAATATCTGTAGAAATTTGCCAGGAGAAATATATATATGATGCAAAACACACAGTGCA 1555
1713 AAGAGTATAAGTACCCCGAAGAGCTGCCGGGACAGTTATACGATGCAAAATACCCAATGCA 1772
1556 AGTGGCAGTTCCGGAGGAAGCCAGCTCTGATGCTGCACTTAAAGAGCACTGTGA 1615
1773 AGTGGCAGTTGGAGGAAAGCCAGCTCTGTATGCTGCACTTCAAGAAAGGACATCTGTA 1832
1616 AAGCCCTGTGGTGCATCTGATTTGGAAGAAATGTGAGACTAAATTTATGCCAGCAGCAG 1675
1833 AGGCCCTTGTGGTGCATCTGATTTGGAAGAAATGTGAGACCAAGTTTCATGCCAGCAGCAG 1892
1676 AAGGCACAAATTTGTGGGCATGACATGTGTGTGCCGGGAGGACAGTGTGTGAAATATGGTG 1735
1893 AGGGTACTCTGTGTGGCAGGACATGTGTGTCTGTGGAGGACAGTGTGTCAAGTACGGTG 1952
1736 ATGAAGGCCCAAGCCCAACCCATGGCCACTGGTGGACTGTGTCTTCTGGTCCCAATGCT 1795
1953 ATGAAGGCCCAAGCCCAACCCATGGCCATTTGGTTCAGATTTGGTCTCCCTGGTCCCTGCT 2012
1796 CCAGGACCTGTGGGGAGGAAATCTCTCACAGAGCCGTCTCTGTACCAATCCAGACCAT 1855
2013 CCAGGACCTGTGGGGAGGAAATCTCTCACAGAGCCGTCTCTGTACCAATCCAGACCAT 2072
1856 CGCATGGAGGAAATTTCTGTAGGGCTCCACTCGCACTCTGAAGCTCTGCAACAGTCA 1915
2073 CTCATGGAGGAAATTTTCAGGGCTCCACAGCCTCTGAAGCTTTGCAACAGTCA 2132
1916 AATGTCCCGGAGCAGTGTGATCTTCGGTGTCTCAGTGTGCGGAGGACACACAGCAGAC 1975
2133 GATGCCCTCTGACACAGTGTGGAATTTCCGTGTCTGCCAGTGTGCCAGTATTAACAGCAAGC 2192
1976 GATTCAGAGGGGGCACTACAGTGGAGCCCTTACACTCAAGTACAGATCAGGACTTAT 2035
2193 GCTTCGAGGGTGGCTCTACAAAGTGAAGCCCTTACACCCAGTTGGA-----AGGACTTAT 2247
2036 GCAAACTCTACTGTATCGCAGAGGATTTGATTTCTTTCTTTTGTCAAAATAAGTCA 2095
2248 GCAAACTCTACTGTATCGCAGAGGATTTGATTTCTTTCTTTTGTCAAAATAAGTCA 2307
2096 AAGATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTTGTATAGATGGGATATGT---- 2151
2308 AAGATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTTGTATAGATGGGATGTGTAGG 2367
2152 ----- 2151
2368 TAATAACGACCGCTTATTCATTAACAAACGCTCTCCAGACAGTCTGTTTGTGCCAGGCA 2427
2152 ----- 2151
2428 TCCCGTGGGGCTAGAAAGCCAGACATATGAAGAGGCCCTTCCCTCCAGTCCAGCC 2487
2152 ----- 2151
2488 CAGGGGGAGAAAGCCAGAGAGCTCATGTGTATGGCAGTCTTTGTGTAGGAGGGTTG 2547
2162 GATGTGCAATGTCTTGTGATCTGATGCTGTTGAAGACGCTCTGTGGGGTGTGTAAACGGGA 2221
2548 GCTGTGACAAATGTCTTGTGATCAGATGCAACAGAGGACTCTGTGGGGTCTGCAAGGGGA 2607
2222 ATAACTCAGCGTGCAGATTTCAAGGGGTCTCTACACCAAGCAGCACCACCAACCACT 2281
2608 ATAACTCAGACTGTGTACGCGCATAGAGGGCTCTACAGTAAAGCACCATTCCCAACCACT 2667

2282 ATTATCACAATGGTCAACATTTCTTGGAGCCCGGAGTATCCGATCTATGAAATGAACG 2341
2668 ACTACACATGTTAAACATTTCTTGGAGCTCGGAGCAATTCACATCTACGAAACAAACA 2727
2342 TCTTACCTCTCTACATTTCTGTGCGCAATGCCCTCAGAAAGGTACTACTCTGAATGGGCAC 2401
2728 TCTTACCTCTCTATATTTCTGTGCGCAATTTCTCTCAAGAGATATTTACCTGAATGGACACT 2787
2402 GGACGCTGTGACTGTGGCCCGCGGTACAAATTTTTCGGGCACTACTTTCGACTACAGACGGT 2461
2788 GGAAGTGTGACTGTGGCTGGCGATACAAGTTTTCGGGTGCCACCTTCAACTACAAACGGT 2847
2462 CCTAATAATGAGCCCGAGAACTTAATGCTACTGTGACCAACCAAGAGACACTGATTGTGG 2521
2848 CTTACAGAGGCTGAGAAATTTAACTCCCTGGACCAACCAATGAAATCTGATTGTGG 2907
2522 AGCTCTGTCTTTCAGGGAAGGAAACCGGGTGTGGCTTGGGAAATATCTCATGCCCTCGCTGG 2581
2908 AGCTTTTGTTCAGGAGGAACTCTGTGTGGCTGTGGGAAATTTTTCATTGCCAAGGTGAG 2967
2582 GGACCGAAGAGCAGCCCTGTCAGCCAGCTACACTTTGGGCGCATCTGTGCGCTCTGAGT 2641
2968 GGGCTAAGAGAGACTCAGCTGCCAGCCAGCTACTCATGGGCGCATCTGTACGCTCTGAGT 3027
2642 GCTCCGTCTCTGCGGAGGGGACAGATGACCGTGAGAGAGGCTGTACAGAGACCTGA 2701
3028 GCTCCGTCTCTGCGAGGGGCAAGATGAATCAAAAGCGGGATGTCTACAGAGACCTGA 3087
2702 AGTTTCAAGTAAATATGTCTTCTGCAATCCCAAGACACGACTGTACGGGCTGTGTC 2761
3088 AGTTTCCGGTGAATGCAATTTCTGCAACCCCAAGACAGACTGTACAGGCTTGTGTC 3147
2762 TTGCAAAAGTATCTCCTGTCTCTCCAGCTGTGTCCGTGGGAACTGGAGTGTGCTGCAGTC 2821
3148 CTTGTAAGGTGTCACTTGTCTTCCAGCTGTCTGTGGGAACTGGAGCGTGTGCAGCC 3207
2822 GGAAGTGTGGGGGGTGGCCAGAGCGCCCGTGTGAGTGTGACACGGCGGGTGTGACTATG 2881
3208 GGACATGCGGTGGGGGACCCAGAGAGCGCCGTGGGTGACACGGAGGGGACACTACC 3267
2882 ACTCGAGGACGCTCCCGGCGACCTGTGTGCTCTCAGCTGTCTCCCTCAGCAGGACGAGCCT 2941
3268 GTGACAGTCTCATCCAGCGGAGCTGTGTCTCACAGCCCGAGCCCTTACCAAGCCT 3327
2942 GCAACTCTCAGAGTCTCCACCTGTATGGAGCGCGGGCCCTTGGGCGAGTGTCTCACACA 3001
3328 GTAACTCTCAGAGTCTCCACCTGTCTGGAGCACTGGGCGCTTGGGCGAGTGTCTCAAGGA 3387
3002 CTTGTGGGAAGGGGTGGAGGAAGCGGCGAGTGGCTGTAAAGACCAACCCCTCGGCCCA 3061
3388 CTTGTGGGAAGGGGTGGAGGAAGAGGACAGTGGCTGTGCAAAAGCACCMAACCCCTCAGCTC 3447
3062 GAGCGCAGTCTGCGCCGACGCTGTGTCACTCCGAGCCCAAGCCAGGATGCAATGAAG 3121
3448 GAGCCAGTGTGTGACAGACACTGTGTCACTTCAAGACCCAGCCCTCGAACCCAGAAA 3507
3122 CTTGTCTGTCTCAGCGCTGCCCAAGCCCAAGAGCTGCAGTGGTGTGTGCGCTGTGT 3181
3508 TCTGCTGTCTCAAGCGCTGCCCAAGCACAAGCAAGCTACAGTGGCTGTGTATCTGCTGTGT 3567
3182 CCAAGTGTCTGTGACATGTGAAGAGGAACAAGAAAGATTTCTTAAATGTGTCTGAA 3241
3568 CCAAGTGTCTGTGACATGTGAGGGGGGACAAACAGAGATTTCTCAGTGTGTGTGAGA 3627
3242 AGTATGTTTCTGGAAGTATCGAGAGCTGGCTCAAGAGAGTGTCAACATTTGCCGAGC 3301
3628 AGTATCATCTCCGGAAGTATCGAGAGCTGGCTTCGAAGAAATGCTTGTGATCTGCCAAGC 3687
3302 CCAGCTGTGAGTGTGAACGTCCTGCGCCCGCTTCCATGCCCCCAGGCAACCCCTCATTTG 3361
3688 CTGACTGTGAGTGTGNAAGCGCTGTGGGCTGTATTTCCCTGCCCCCAACACCTCCGTTG 3747
3362 CTGCTGCGGACCCCTCGAGGGGCAAGTGTGTTTGGCTCTCACCTGTGTCTCAGTGCACGGCCA 3421

Query Match		30.5%;	Score 1120.8;	DB 3;	Length 4335;
Best Local Similarity		61.5%;	Pred. No. 2e-300;		
Matches 2015;		Conservative	0;	Mismatches 1187;	Indels 72; Gaps 11;
Qy	433	TGTTTTCTATCAAGGCTTTTGGGATCACAGAAATCTCTCAGTGGCCCTTTCAACCTGC	492	1428	CGAGTATCTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCTTGTCTGATCAGCCAAA
Db	898	TGTTTGATCAGGATTTATCAGAAATGACAGCTCGTCTCTGTGGCTGTGTCTACGTGT	957	1912	ACAGTATCTAAAGAAATTTCTCAGTACACCGCAGGCTGGCTGTCTGTGTGATGAGCCCAA
Qy	493	CRAAGGCTTGTGAGGATGATACGAAACAGAGAGGCGAGATTACTTCTTAAGGCGCACTTCCT	552	1488	GCCTGTGAAGGAATACAAGTATCTCAGAAATTTCCAGAGAAATTCAGAGAGATTAATATCATGCAACAC
Db	958	GCCGGCTTGTGAGTCTGATAAGGACCGGAGACATGATTTCTCTCATCTCACCGCTACCT	1017	1972	GCAACGGGACAGTATAAATATCCGGACAAACTCCCGGGACAGATTATATGATGCCGACAT
Qy	553	TCAACACTCTCATGGAATCTGGGAGAGCTGGCCAAAGGCGAGCTGGCAATCCCAAGTACTG	612	1548	ACAGTCAAGTGGCAGTTCGGAGAGAAAGCCAAAGCTCTGTCATCTGAGCTTTTAAAAAGGA
Db	1018	CAGCTGTGGCCCGAGGAACACAACTATAGCTCGCTGCGAGGCCACCATCTCACGCTCCTG	1077	2032	GCAATGTAGTGGCAATTTGGAGCCAAAGCCAAAGCTGTGACGCTTTGGGGTTATGAGGA
Qy	613	TACAAGAGATCCACAGGCCCTATGCTCTCTGGGGCCAGTGAGTCTCTGTGAGCTCAAGG	672	1608	CATCTGTAAAGCCCTGTGTGCCATCGTATGTGAAGAAATGTGAGACTAAATTTATGCC
Db	1078	TACAAAGGACAGCAGAGAGAGGGTCAGATGGTA CCAAGACTTACCTTGGCTCCAGCGG	1137	2092	CATTTGCAAGTCACTCTGTGTGCCACCGCGTGGGCCACAGATGTGAACCAAGTTTCATGCC
Qy	673	ACAT-----GGGAGCTGGCACATCAACCCCTGCGACAGCGAGCCTTCGCTCGG- 721		1668	AGCAGAGAGGACACAATTTGTGGGCATGACATGTGTGCGCGGGAGGACAGTGTGTGAA
Db	1138	ACATATCCTGGTCACTCCCAAGTCACTCTCCCTGCTCCAGAGCCAAAGAGCCAGAG	1197	2152	TGCAGCAGAGGGACCGCTTGTGGCTTGTGATATGTGTGTGTGGCAAGGCCAGTGTGTAAA
Qy	722	-----GACTGCCAAAAAGCAGCATTTCTGTGGAAGACGCAAGAAATACAT- 767		1728	ATATGTGTATGAAGGCCCAAGCCCACTATGGCCCACTGTGTCCGAGCTTGTCTTTGGTC
Db	1198	TACAGCCATCGAAGGTGGCAAAAGCGCATTTTGTGGACGACGCAAGAAATGTATATGC	1257	2212	GCTCGGAGAACTAGGGCCCGGCCCATCCATGGCCAGTGTCTGCTGTGCGAAGTGTGTC
Qy	768	GCCCCAGCTCCCAAGGAGACCTCTTCATCTTGCAGATGATATAGTCTTGTCTAGC	827	1788	CCCATGCTCAGAGACCTGCGAGGGGAGTATCTCATAGAGTGCCTCTGCAACCAACC
Db	1258	CCCAAGCCTCTGCGAGAGGACCCCTATCTACGCTTTGACGAATATGGAGGCGACAGGGCG	1317	2272	AGAAATTTCTCGCACTTGGCGGAGAGTCAAGTTCCAGGAGAGACACTGTAGTAACCC
Qy	828	GCATAAGCGCTCTTCTGAGTCCCATGAATGAAGACTCAACGTGGAGACCTTGGT	887	1848	CAAGCCATCGCATGGAGGGAATTTGTGAGGGCTCCACTCGCACTCTGAGCTCTGCAA
Db	1318	GCCACAGGCTGAGCTGGAAGTCAAA-----ATGGTCTCAATGTGGAGACCTTGT	1371	2332	CAAGCCTCAGTATGTGTGGCAAGTACTGTCCAGGATCTAGCCGCACTTACAAGCTGTGCAA
Qy	888	GGTGGTGCACAAAAGATGATGCAAAACCATGGCCATGAAAATATCACCACTTACGTGT	947	1908	CAGTCAAGAAATGTCCCGCGGACAGTGTGACTTCCGTGTGCTCAGTGTGCCGAGACAA
Db	1372	TGTGGCAGATGCCAAGATGTGGAGAGCAGCGCAAGGATGACGTCAACCAGTACATTTCT	1431	2392	CATTAACCTTGTCTCGAGATAGCTTGGATTTCCGTGCCCAACAGTGTGCAAGTATAA
Qy	948	CAGGATCTCAACATGGTATCTGCTTTATTCAAAGATGGAAACAAATGAGGAGAAACATCAA	1007	1968	CAGCAGACGATTCAGAGGGCGGCATCAAGTGGAAAGCCTTACACTCAAGTGAAGATCA
Db	1432	CACAGTCATGAACATGGTTTCTAGCCTGTTCAAAGATGGGACCATTTGGAAGCGACATAA	1491	2452	CAACAGCCCTTCCGTGATGGTTGTACCGTTGGAACCCCTACACAAAAGTTGAAGAGGA
Qy	1008	CATTGCAATGTAGTCTGATCTTCTAGAGATGAACAGCCAGGACTGGTATAGTCA	1067	2028	GGACTTATGCAAACTTACTGTATCCAGAGGATAGCCGTAATGTTTGTATAGATGGAT
Db	1492	CATTGTGTTGTGAGCCTTAATCTGCTGGAAGAAGAACCTGAAAGGACTGCTGATCAATCA	1551	2512	AGATCGATGTAACTCTACTGCAAGCGGAGAACTTTGAGTGTCTTTGCCATGTCTGG
Qy	1068	CCACGACAGACCAACCTTAAGTACTTCTGCCAGTGGCAGTCTGGATTGATGGGGAAGA	1127	2088	TAAAGTCAAGATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTTGTATAGATGGAT
Db	1552	CCATGACAGACCACTCTAACAAGCTTCTGTCAATGGCAGTCACTTGTGGGAAGAA	1611	2572	CAAGGTGAAGATGGAGCCCTTGTTCGCCACACAGAAATGACGTCTGCATCATGGGAT
Qy	1128	TGGGACTGTGATGACACGCCATCTTACTGACTGTGTGATATATATGTTCTCGGAAGAA	1187	2148	ATGTGAGAGAGTGGATGTGACAAATGCTCTGATCTGATGTGTTGAAGACGCTGTGTGG
Db	1612	TGGCAAGAGACAGCAGCAGCCATCTCTCAAGGATTTGACATTTGTTCTCTGGAAGAA	1671	2632	TTGTGAACCTGGTGGGATGTGACCATGAGCTTGGCTCTAAGGCGGTTTCTGATGCATGTGG
Qy	1188	TGAGCCCTGTGACACTTTGGGATTTGCACCCATAGTGGAAATGTGTAGTAAATATCGCAG	1247	2208	GGTGTGTAACGGGAATAACTCAGCCCTGACGATTCACAGGGGTCTCTACACCAAGACCA
Db	1672	CGAACATGTGACACTAGATTTGCTCTTATCAGTGGCATGTGCAATGATGACGAAAG	1731	2692	TGTTTCAAGGCGATAACTCAACTTGTCAAGTTTTATAAGGGCTGTACTCTCAGTCAGCA
Qy	1248	CTGCAGATTAATGAAGATAACAGTCTTGGACTTGGCTCTTCCATTTGCCCATGAGTCTGG	1307	2268	CCACACCAACCAAGTATATCATGTGTCAACCTCTCTTGGAGCCCGGAGTATCCGCAT
Db	1732	CTGTACCATCAATGAGACACAGACTTGGCTTGGCTTCCACCATTTGGCATGAGTCAGS	1791	2752	CAAGCAATGAGTATATACCTGTGTGTACCATCCCGCTGGGGCCCGAAGCATTTGAGAT
Qy	1308	ACACACTTTGGCATGATGAGAGAGGGAACATGTGTAAAGTTCGAGGGGCAA	1367	2328	CTATGAATGAAGCTCTTACCTCTCAATTTCTGTGCGCAATGCGCTCAGAGAGTACTA
Db	1792	GCACAACTTTGGCATGTTTCTATGATGTGTGAAGCAATCCCTGCGAGGAAGCAGAGGCAA	1851	2812	TCAGGAGTTGCACTTTCTTCAGCTATCTTGTCTGTTCGAAGCCTCAGTCAAAAGTATTA
Qy	1368	CATCATGTCCCTTACATGGCAGGACGATGAGTCTTCTCTGTGTCACTTGTGAGCCG	1427	2388	CCTGAATGGGCACTGACCGTGTGAGCTGGCCCGGCTGGTACAAATTTTTCGGGCACTACTTT
Db	1852	CATCATGTCCACCACTGACTGGAACCAATGGGGGTGTTTTTCAATGTGTGCTGCGACCGG	1911	2872	CCTCACAGGGGGTGGAGATCGACTGGCCAGGGGACTTCACTTTCGCAAGGCAACGCTT
Qy				2448	CGACTACAGACCGCTCTAATAGCCCGGAGAACTTAAATCGCTTACCTGGAACCAACCA
Db				2507	CGACTACAGACCGCTCTAATAGCCCGGAGAACTTAAATCGCTTACCTGGAACCAACCA
Qy				2508	GACACTGATTTGTGGAGCTGCTGTTTTCAGGGAAAGGAAACCGGGGTGTTGCTTGGGAATATC

Db 2992 GACGCTGCTTTGAAATCTTGACACAGGCAAGAAATCCGGGATCGCTTGGAGTATGC 3051
Qy 2568 CATGCTCGCTTGGGAGCGAGAGCAGCCCGCTGCCAG---CCAGCTACACTTGGGC 2624
Db 3052 ACTTCCCAAGGTCATGAATGTAACTCAGCAGGCCAACCAAGAGATACCAACCAACCTGGCG 3111
Qy 2625 CATCGTCGGCTGAGTCTCGTCTCGTCTCGGAGGGGACAGATGACCGTGAGAGGG 2684
Db 3112 CACAGTGAGTCTGACTGCTCAGTCACCTGCGGTGGAGGTATATAAGCATCAAGGCCAT 3171
Qy 2685 CTGCTACAGAGACCTGAAATTTCAAGTAAATATGCTTCTTGCAATCCCAAGACACGACC 2744
Db 3172 TTGCTTACAGAGATCAACACCAAGTCAATTCCTCATTTCTGCACTGTAAGAACCAAGCC 3231
Qy 2745 TGTACGGGGTGTGCTTGTCAAGATATGTCCTGTCTCTCCAGCTGCTGTCGGTGGGAA 2804
Db 3232 AGCAACAGAGCCCAAGATATGCAAGCTTTCTCTGCGCAGCTATTGTTGCCAGGGGA 3291
Qy 2805 CTGGAGTCCCTGAGTCGAGCGTGTGGGGGGTCCCGAGAGCCGCCCGTGCAGTGCAC 2864
Db 3292 GTGGAGCGATGACGAAATGCTGTGCGCGGCGCCAGCAGAGCCGCAAGATACGGTGTGT 3351
Qy 2865 ACGCGGGTGCACCT---ATGACTCGGAGCCAGTCCCGGCCAGCTGTGCTCCTCAGCCTGC 2921
Db 3352 CGAAGAAGCCCTTCCAGAAGAGAGGCGGTGCTGCACCTCTCTGCCAGTGCAC 3411
Qy 2922 TCCCTCCAGCAGCAGGCTGCACTCTCAGAGTGTCCACCTGATGAGAGGCGCGGCC 2981
Db 3412 GCCACTCAGGTTCAAGTCTGCAACAGCATGCTGCTGCGCTCCAGAGTGGAGGCCAGGCC 3471
Qy 2982 CTGGCAGAGTGTCTCACACCTGTGGNAGGGGTGGAGGAGCGGGCAGTGGCTGTAA 3041
Db 3472 GTGGTCTCAGTGTTCAGAGCTGTGGAGGCGGTGAGGCGGTGAGGCTCTTTGTAA 3531
Qy 3042 GAGCACCAACCCCTCGGCGAGAGCAGCTGTGCTCCGACGCTGTCTGCACCTCCGAGCC 3101
Db 3532 AGCCCTGC-----AGCGGAGACACTCCCGAGAGCCTGTGCTCCAGCAGGCC 3579
Qy 3102 CAAGCCAGAGATGATGAAGCTGTCTGCTTCCAGGCTGCGCCACAGCCCAAGAGCTGCA 3161
Db 3580 CAGACCCAGGACACAGGAGGGTGTGTGCTGGGAGCGTGCCCAAAAAACAACAGGCTCCA 3639
Qy 3162 GTGGCTGTGTCGGCTGTGCTCCAGTGTCTGTGATGTGAAGAGGAGGAGGAGGAGGAG 3221
Db 3640 GTGGATGTCTGATGAGTGTGTTCTGCAACCTGTGCTGCTGGGTGGGTGAGGAAGAG 3699
Qy 3222 ATTCTTAAATGTGTGAAAAGTATGTTTCTGAAAAGTATCGAGAGCTGGCCTCAAAGAA 3281
Db 3700 AGAATCTGAATGCGTTGAGAAGACCTTGACGGGGAACTAATAACATTCGCGAGCGGAG 3759
Qy 3282 GTGCTACATTTGCGGAGGCCA-GCCTGGAGCTGGAAGCTGCTGCGCGCCCGCTTCAT 3340
Db 3760 ATGCGGCAACATTAAGAGGCCAAGCCTGGAATCTGGAAGAGCCTGCAACCAAGACTT 3819
Qy 3341 GCCCAGGACCCCAATTTGCTGCTGCGGACCCCTCGAGGGGAGCTGTTTGCCTCAC 3400
Db 3820 GCCAGTGTAC-----AGCATGGCGTGGCGAGCTGTGATTCATCAC 3861
Qy 3401 CTTGTGCTCAGTGCAAGCGGCTGCTGGGGGAGGGTTTCAGACGAGGTCGTCGAG-TGC 3459
Db 3862 CATGCAACAGTGCAAGTAACTCTGCGGGGAGGAGTCCAGACTCGCTCTGTCCAGCTGC 3921
Qy 3460 CTGGCTGGGGGGCGCGCGCTCAGGCTGCTCTGTCGACACAGAGCTTGGCGCTCCCTG 3519
Db 3922 ATGCAGCAAGGGCGGCTTCTCCTCAAGTTGTCTGCTCCATCAGAGCCTCCGCTGCTCAG 3981
Qy 3520 GCTGCAACACTCACTTCTGCCCCATTGCGAGAGAGAGATG---CCTTCTGCAAGAC 3576
Db 3982 GCTGTAAACCAACACTTCTGTCAGCTCTGAAAAGAAAGATGACCCATCTTGTGGTGGAT 4041
Qy 3577 TACTTCCACTGGTGTGTACCTGGTACCCCAAGCAGCGGATGTGACGACCAAGTTCTACGCG 3636

Db 4042 TTCTTACGTGGTGTACCTGGTCCCTCAGCACGAGTCTGTCAACCACAAAATTTATGGA 4101
Qy 3637 AAGCAGTCTGCAAGACTTGTCTTAAGTCCAACT 3670
Db 4102 AAGCAGTCTGTAGATCATGCACAAGGAAGACT 4135

RESULT 3
BM808410
LOCUS
DEFINITION BM808410 1092 bp mRNA linear EST 05-MAR-2002
5', mRNA sequence.
AGENCY 6604069 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5591626
ACCESSION BM808410
VERSION BM808410.1 GI:19125233
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1092)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12367 row: n column: 11
High quality sequence stop: 728.
Location/Qualifiers
1. 1092
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5591626"
/lab_host="DH10B"
/clone_lib="NIH_MGC_125"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

FEATURES
source
1. 1092
Query Match 22.0%; Score 808.4; DB 4; Length 1092;
Best Local Similarity 97.2%; Pred. No. 1.8e-213;
Matches 855; Conservative 0; Mismatches 21; Indels 4; Gaps 3;

ORIGIN
Qy 2237 CGATTTCAGGGTCTCTACACCAAGCACCACCAACAGTATTATCATGGTCA 2296
Db 19 CGATTTCAGGGTCTCTACACCAAGCACCACCAACAGTATTATCATGGTCA 78
Qy 2297 CCATTTCCTTCTGAGCCCGGAGTATCCGCATCTATGAATGAACGTCTTACCTCTTCA 2356
Db 79 CCATTTCCTTCTGAGCCCGGAGTATCCGCATCTATGAATGAACGTCTTACCTCTTCA 138
Qy 2357 TTTCTGTGCGCAATCCCTCAGAGGTAATACTGAAATGGGCACTGGACCGGTGGACTGCG 2416
Db 139 TTTCTGTGCGCAATCCCTCAGAGGTAATACTGAAATGGGCACTGGACCGGTGGACTGCG 198
Qy 2417 CCGGCCGCTACAAATTTTGGGCACTACTTTTCGACTACAGAGGTCCTATATGAGCCCG 2476
Db 199 CCGGCCGCTACAAATTTTGGGCACTACTTTTCGACTACAGAGGTCCTATATGAGCCCG 258

```
QY 2477 AGAAGCTTAATCGTACTGAGCAACCAACGAGACACTGATTGTGGAGCTGCTGTTTCAGG 2536
Db 259 AGAAGCTTAATCGTACTGAGCAACCAACGAGACACTGATTGTGGAGCTGCTGTTTCAGG 318
QY 2537 GAAGGAACCCGGGTGTGCTGGGAATATCCATGCTCGCTTGGGGACCGAGAGCAGC 2596
Db 319 GAAGGAACCCGGGTGTGCTGGGAATATCCATGCTCGCTTGGGGACCGAGAGCAGC 378
QY 2597 CCCTGCCCAGGCCAGCTACACTTGGGCCATCGTGGCTCTGAGTGTCTCGTGTCTCGG 2656
Db 379 CCCTGCCCAGGCCAGCTACACTTGGGCCATCGTGGCTCTGAGTGTCTCGTGTCTCGG 438
QY 2657 GAGGGGACAGATGACCGTGGAGAGGGCTGTACAGAGACCTGAAGTTTCAAGTAATA 2716
Db 439 GAGGGGACAGATGACCGTGGAGAGGGCTGTACAGAGACCTGAAGTTTCAAGTAATA 498
QY 2717 TGTCTCTTCTGCAATCCCAAGACACGACCTGTACGGGGCTGTGCTTCAAGTAATCTG 2776
Db 499 TGTCTCTTCTGCAATCCCAAGACACGACCTGTACGGGGCTGTGCTTCAAGTAATCTG 558
QY 2777 CTGTCTCTCCAGCTGGTCCGTGGGAACCTGGAGTGCCTGCGAGTCGGAGCTGTGGCGGG 2836
Db 559 CTGTCTCTCCAGCTGGTCCGTGGGAACCTGGAGTGCCTGCGAGTCGGAGCTGTGGCGGG 618
QY 2837 GTGCCAGAGCCGCCCTGTCAGTGTGACACGGGGGTGCACTATGACTCGAGCCAGTCC 2896
Db 619 GTGCCAGAGCCGCCCTGTCAGTGTGACACGGGGGTGCACTATGACTCGAGCCAGTCC 678
QY 2897 CGGCGAGCTGTGCTCAGCTGCTGCTCCCTCCAGCAGGCGAGCGCTGCACTCTCAGAGCT 2956
Db 679 CGGCGAGCTGTGCTCAGCTGCTGCTCCCTCCAGCAGGCGAGCGCTGCACTCTCAGAGCT 738
QY 2957 GCCCAGCTGTCATGGAGCGCGGGG-CCCTGGGCGAGTGTCTCACACACCTGTGGGAA--GG 3013
Db 739 GCCCAGCTGTCATGGAGCGCGGGGCGGCGCTGCGGCGAGTGTCTCACACACCTGTGGGAA 798
QY 3014 GTGGAGGAAGCGGCGAGTGGCTGTGAAGACCAACCCCTCGCGCAGAGCGAGCTGC 3073
Db 799 GTGGAGGAAGCGGCGAATGGCCCTGTAAACCAACCACTTCGGGCGAGACCCCACTGC 858
QY 3074 TG-CCGAGCGTGTGCACTCCGAGCCCAAGCCAGCA 3112
Db 859 TGCCCAAGAGCTGTGTGCACTCCGAGCCCAAGCCCAAGA 898

RESULT 4
AK036137
LOCUS
DEFINITION
Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:9630038L21 product:hypothetical
Metalloprotease (ADAM type)/reprolysin (M12B) family containing
protein, full insert sequence.
ACCESSION
AK036137
VERSION
AK036137.1 GI:26085046
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED
11042159
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
20499374
PUBMED
11042159
```

```
REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Hazada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, O.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
11076861
PUBMED
11076861
REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6
(bases 1 to 4593)
REFERENCE
AUTHORS
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Koya, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
FEATURES
source
location/Qualifiers
1..4593
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM,DB:9630038L21"
/db_xref="taxon:10090"
/clone="9630038L21"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days neonate"
misc_feature
1..4593
/note="hypothetical Metalloprotease (ADAM type)/reprolysin
(M12B) family containing protein (InterPro|IPR001590,
evidence: InterPro)"
ORIGIN
Query Match 21.7%; Score 795.8; DB 3; Length 4593;
Best Local Similarity 57.1%; Pred. No. ie-209;
Matches 1882; Conservative 0; Mismatches 1117; Indels 296; Gaps 12;
```

QY 478 GCCCTTCAACCTGCCAAGGCTTGTTCAGGCATGATACGAAACAGAGAGGCGAGATTACTTTC 537
Db 1 GTGTGTCTACGTGTGCCGGCTTGTTCAGGTCTGTAAAGGACCCGAGACAATGAATTCCTTC 60
QY 538 CTAAGGCGCACTTCTCTCAACCTCTCATCGAAACTCGGCAGAGCTGCCCAAGCGAGCTCG 597
Db 61 ATCTCACCGCTACCTCAGCTGTGTGGCCAGGAAACAACTATAGCTCGCTGCAGGCCAC 120
QY 598 CCATCCCAAGTACTGTACAAAGAGATCCACAGAGCCCCATGTCTCTGGGGCCAGTAGGTTC 657
Db 121 CATCTCTCAGGTCTGTGTACAAAGGACAGCAGAGAGAGGGTTCAGATGTGTACCAAGACTAC 180
QY 658 CTGCTGACCTCAAGGACATGGGAG-----CTGGCACATCAACCCCTGCACAGCAG 707
Db 181 CTGTGGCTCCAGGGAGCATCTCTGTGTCACTCCCAAGTCAACATCTCCCTGCTGCCCCAG 240
QY 708 CGACCTTTCGCTG-----GGACTGCCACAAAAGCAGCATTTCTGTGGAGAGCGC 756
Db 241 AGCCAAGAGCCAGAGTACAGCCATCGAAGGTGSCAAAGCGGCAATTTGTGTGACGACGC 300
QY 757 AAGAAATACATG-----CCCCAGCTCCCAAGAGACCTCTTCATCTTGGCCAGATGAGTA 812
Db 301 AAGAAATGTATATGCCCCAAGCCTCTCTGCAGAGGACGCTTATCTACGCTTTGACGAATA 360
QY 813 TAAGTCTTGTCTTACGGCATAAGCGCTCTCTCTGAGGTCCCATAGAAATGAAGAACTGAA 872
Db 361 TGGAGGCACAGGGGCCCGAGCGGTACGCTGGAAGTCAAAA-----ATGGTCTCAA 414
QY 873 CGTGGAGACTTTGGTGGTGTGCGACAAAAGATGATGCAAAACCATGGCCATGAAATAAT 932
Db 415 TGTGGAGACCTTGTGTGGCAGATGCCAAGATGTGTGGAGAGCAGCGCAAGGATGACGT 474
QY 933 CACCACTACGTGCTCAGATACTCAATGTGTATCTGCTTTATTTCAAGATGGAACAAT 992
Db 475 CACCAAGTACATCTCTCACAGTCAATGAACATGTGTTCTAGCCCTGTTCAAAGATGGGACCAT 534
QY 993 AGGAGGAAATCAACATTGCAATTTAGTGTCTGATTTCTCTAGAAAGATGAAC----- 1045
Db 535 TGGAGCGACATTAACATTTGTGGTGTGAGCCTAATTTCTGTGGAGAGAAACCTCTACA 594
QY 1046 ----- 1045
Db 595 AAGGCTCAACCTTCTCTCTCGGCACAAATCATCACTTTTAATCCCTGCGACCTTTCCCA 654
QY 1046 -AGCCAGGACTGATAGTACCAAGCAGACCACTTAAGTAGCTTCTGCCAGTGG 1104
Db 655 TAGGAAGGACTGCTGATCAATCACCATGAGAGCCAGTCTCTAAAAGCTTCTGTCAAGTGG 714
QY 1105 CAGTCTGGAATTGATGGGAAAGATGGGACTCGTCTATGACCGCCATCTTACTGACTGGT 1164
Db 715 CAGTCAGCTCTTGTGGGAAGAAATGGCAAGAGACACGACCGCCATCTCTCTCACAGGA 774
QY 1165 CTGGATATATGTTCTGGGAAGATGAGCCCTGTGACACTTTGGGANTTTGCCACCAATAGT 1224
Db 775 TTTGACATTTGTTCTCTGGGAAGCAACCAATGTGACACACTAGGATTTGCTCCTATCAGT 834
QY 1225 GGAATGTGTAGTAAATATATCGCAGCTGACGATTTAATGAAGATACAGGCTTTGGACTGGCC 1284
Db 835 GGCATGTGCGTAAGTACCGAAGCTGTACCATTAATGAAGACACAGGACTTGGCCCTGGCC 894
QY 1285 TTTACCATTTGCCCATGTCTGGACAACTTTTGGCATGATTCATGATGGAGAAGGGAAAC 1344
Db 895 TTTCAAAATTTGCGCATGAGTCAGGCGACAACTTTTGGCATGTTTCATGATGGTGAAGGCAAT 954
QY 1345 ATGTGTAAAAGTCCGAGGGGACATCATGTCCCTACATTTGGCAGGAGCGAATGGAGTC 1404
Db 955 CCGTCAGGAAAGCAGAAAGGCAACATCATGTCAACCACTGACTGGAACAAATGGGGTG 1014
QY 1405 TTTCTCTGTCTACCTTGCAGCGCCAGTATCTTACAAATTTCTAAGCAGCGCTCAAGGT 1464
Db 1015 TTTTTCATGCTCGTCTGAGCCGACAGTATCTTAAGAAATTCCT----- 1058

QY 1465 ATCTGCCTTGTGCTGATCAGCCAAAGCCTGTGAAGGNATACAAGTATCTCTGAGNAATTGCCA 1524
Db 1059 ----- 1058
QY 1525 GGAGAAATTATATGATGCAAAACACACAGGTGCAAGTGGCAGTTTCGGAGAGAAAGCCCAAGCTC 1584
Db 1059 ----- 1058
QY 1585 TGCATGTGCGGACTTTTAAAAAGGACATCTGTAAAAGCCCTGTGTGGTCCATCTGTATTTGGAAG 1644
Db 1059 -----CAGGACATTTGCAAGTCACTCTGTGTGCCACCGCGTGGGCCAC 1100
QY 1645 AAATCTGAGACTAAATTTATGCCAGCAGCAAGGCAAAATTTGTGGGCATGACATGTGG 1704
Db 1101 AGATGTGAACAACAGTTTCATGCTGCAGCAGAGGGGACCGCTTGTG-----GCTTGATGG 1155
QY 1705 TGCCGGGAGGACAGTGTGTGAAATATATGATGAAGGCCCCCAAGCCCAACCCATCGGCCAC 1764
Db 1156 TGTGGCAAGGCCAGTGTGTAAAGCTCGGAGAACTTAGGGCCCGCGCCCATCCATGGCCAG 1215
QY 1765 TGGTCGGACTGCTCTTCTTGTGCCATGTCTCCAGGACCTCGGAGGGGGAGTATCTCAT 1824
Db 1216 TGGTCTGCTGCTGGAAGTGGTCAAGATTTCTCGCACTTTCGGCGGAGGAGTCAAGTTC 1275
QY 1825 AGGAGTCCCTCTGCACCAACCCCAAGCCATCGCATGGAGGAAAGTTCTGTGAGGGCTCC 1884
Db 1276 CAGGAGAGACACTGTAGTAACCCCAAGCTCAGTATGTGTGCAAGTACTTCCAGGATCT 1335
QY 1885 ACTCGCACTCTGAAGCTCTGCAACAGCTCAGAAATGTCTCCCGGGACAGTGTGTGACTTCGGT 1944
Db 1336 AGCCGCACTCTAAGCTGTGCAACATTAAACCTTGTCTCGAGAAATAGCTTGGATTTCCGT 1395
QY 1945 GCTGCTCAGTGTGCGAGCACAAACAGCAGAGCATTCAGAGGGCGGCACACTACAAGTGGAG 2004
Db 1396 GCCCAACAGTGTGAGAGTATAACAACAGCCCTTCGCTGGATGGTGTACCGTGGGAAA 1455
QY 2005 CCTTACACTCAAGTAGAAGATCAGGACTTATGCAAACTCTACTGTATCGCAGAAAGATT 2064
Db 1456 CCTTACACAAAGTTGAAGAGGAGATCGATGTATAACTTACTGCAAGCGGAGAACTTT 1515
QY 2065 GATTTCTTTTCTTTTCTTTGTCAAAATAAAGTCAAAAGTGGGACTCCATGCTCGGAGGATAGC 2124
Db 1516 GAGTTTTTCTTTGCGCATGTCTGCAAGGTGAAGATGGGACGCTTGTTCGCCACACAGA 1575
QY 2125 CGTAAATGTTCTATAGATGGGATATGTGAGAGATGTGATGTGACAAATGCTTTGGATCT 2184
Db 1576 AATGAGTCTGCATCGATGGGANTTTGGAACCTGGTGGGATGTGACCATGAGCTTGGCTCT 1635
QY 2185 GATGCTGTGAAGAGCTCTGTGGGGTGTGTAAACGGGAAATACTCAGCCTGCACGATTCAC 2244
Db 1636 AAGCGGTTTTCTGATGATGTGGTGTGTTGCAAGCGGATACTCAACTTGCAAGTTTTAT 1695
QY 2245 AGGGGTCTTACACAAAGCACACACACCAACAGATATATACATGCTTACCATTTCT 2304
Db 1696 AAGGGCTGTACCTCAGTCAGCAAAAACAAATGAATATTACCTGTGTGTCACCATTCGCC 1755
QY 2305 TCTGAGCCCGGAGTATCCGCTATGTGAATGAACGCTCTACTCTCTACATCTTCTGTG 2364
Db 1756 GCTGGGGCCCAAGCATTTGAGATTCAGAGTTGAGCTTTCTTCAGCTATCTTGTCTGT 1815
QY 2365 CGCAATGCCCTCAGAAAGTACTACTGAAATGGGCACTGGACCGTGGACTGCCCGGGCGG 2424
Db 1816 CGAAGCCTCAGTCAAAAGTATTACCTCAAGGGGCTGGAGCATCGACTGCGCAGGGGAC 1875
QY 2425 TACAAATTTTCGGGCACTTCTCGACTACAGAGGTCCTATATATGAGCCCGAGAACTTA 2484
Db 1876 TTCACCTTCGCGAGGACACAGTTTGAATACAGCGTTCCTTTAACCGCCCTGGAACGACTG 1935
QY 2485 ATCGCTACTGACCAACCAACAGACACATGATTTGTGGAGCTGCTGTTTTCAGGGAGGAAC 2544
Db 1936 TATGACACGAGGACCCAGCAATGAGACGCTGTCTTTGAAATTTCTGACACAGGCAAGAT 1995
QY 2545 CCGGGTGTGCTGGGAATACTCCATGCTCGCTTGGGGACCGGAGACCGGACCGCCCTGCC 2604


```
QY 514 CGACAGAGAGGAGGAGTACTTCTTAAGCCACTTCTTCAACCTCTCATGGAACTC 573
Db 301 CGAACAGAGAGGAGGAGTACTTCTTAAGCCACTTCTTCAACCTCTCATGGAACTC 360
QY 574 GGCAGAGCTGCCCAAGGAGCTGCCCATCCCACTACTGTACAAGAGATCCACAGAGCC 633
Db 361 GGCAGAGCTGCCCAAGGAGCTGCCCATCCCACTACTGTACAAGAGATCCACAGAGCC 420
QY 634 CATGCTCTGGGCGCAGTGAAGTCTGTGTGAAGTCTCAAGGACATGGAGCTGGCAATCAA 693
Db 421 CATGCTCTGGGCGCAGTGAAGTCTGTGTGAAGTCTCAAGGACATGGAGCTGGCAATCAA 480
QY 694 CCCTGTCACAGGAGGAGCTTGGCTGGGAGTGGCAACAAAGAGAGATTTCTGTGAAGA 753
Db 481 CCCTGTCACAGGAGGAGCTTGGCTGGGAGTGGCAACAAAGAGAGATTTCTGTGAAGA 540
QY 754 CGCAAGAAATACATGCCCCAGCTCCCAAGGAAGACCTCTTCAATCTTGGCAGATGATAT 813
Db 541 CGCAAGAAATACATGCCCCAGCTCCCAAGGAAGACCTCTTCAATCTTGGCAGATGATAT 600
QY 814 AGTCTTGTCTTACGAGATTAAGGCTCTCTTCTGAGGCTCCATAGAAATGAAGATGAAC 873
Db 601 AGTCTTGTCTTACGAGATTAAGGCTCTCTTCTGAGGCTCCATAGAAATGAAGATGAAC 660
QY 874 GTGGAGACCTTGTGTGTGTGCACAAAGAGATGATGCAAAA--CCATGGCCATGAATAT 932
Db 661 GTGGAGACCTTGTGTGTGTGCACAAAGAGATGATGCAAAA--CCATGGCCATGAATAT 720
QY 933 CA-CCACCTACGTCTCAGTACTCAACATGGTA--TCTGCTTTATTTCAAGATGGAAAC 989
Db 721 CACCCCTTACGGCTTACCACTACTCAACATGGTAATCTGCTTTTATTTCAAGAGGGAAC 780
QY 990 AATA-GGAGGAACATCAACATTCGATTTGTA--GGTCTGATCTTCTAGAGATGAACA 1046
Db 781 AATAGGGGGGAACCTTCAACTTGGCAATTTGGACGGGCTGATTCCTTTCAAGAGATGAACA 840
QY 1047 GCCA--GGACTGGTGTATAGTCACCAACGAG--ACCACACCTTAAAGTACTTCTCCAGT 1102
Db 841 GCCAGGACTGGGGGATAGTTCCCTCCCGGACCCCTTATAGTACTTCTGCCAG 900
QY 1103 GG 1104
Db 901 GG 902

RESULT 6
BM850160
LOCUS
DEFINITION
K-EST0130714 S12SNU216 Homo sapiens cDNA clone S12SNU216-81-C06 5',
mRNA sequence.
ACCESSION
BM850160
VERSION
BM850160.1 GI:19206559
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 625)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kr.ibm.re.kr
Plate: 81 row: C column: 06
High quality sequence stop: 625.
```

FEATURES

source

Location/Qualifiers

```
1..625
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S12SNU216-81-C06"
/sex="P"
/tissue_type="Lymph node"
/cell_type="Epithelial"
/lab_host="Top10P"
/clone_lib="S12SNU216"
/note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid phosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10P by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
```

ORIGIN

```
Query Match      17, 0%; Score 623.4; DB 4; Length 625;
Best Local Similarity 99.8%; Pred. No. 6.3e-162;
Matches 624; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 943 GTGCTCAGATPACTCAACATGGTATCTGCTTTATTAAAGATGGAACAATAGGAGGAAC 1002
Db 1 GTGCTCAGATPACTCAACATGGTATCTGCTTTATTAAAGATGGAACAATAGGAGGAAC 60

QY 1003 ATCAACATTGCAATTTAGTCTGATTTCTTAGAAGATGAACGCCAGGACTGGTGATA 1062
Db 61 ATCAACATTGCAATTTAGTCTGATTTCTTAGAAGATGAACGCCAGGACTGGTGATA 120

QY 1063 AGTCACACCGCAGACCAACACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGG 1122
Db 121 AGTCACACCGCAGACCAACACCTTAAGTAGCTTCTGCCAGTGGCAGTCTGGATTGATGGG 180

QY 1123 AAAGATGGGACTGCTATGACACCGCATCTTACTGATCTGGTCTGGATATATGTCCTGG 1182
Db 181 AAAGATGGGACTGCTATGACACCGCATCTTACTGATCTGGTCTGGATATATGTCCTGG 240

QY 1183 AAGAAATGAGCCCTGTGACACTTTGGGATTTGCACCCATNAGTGAATGCTGTAGTAATAT 1242
Db 241 AAGAAATGAGCCCTGTGACACTTTGGGATTTGCACCCATNAGTGAATGCTGTAGTAATAT 300

QY 1243 CGCAGCTGCAGATTAATGAAGATACAGGCTCTTGGACTGGCTTTCAACATTTGCCCATGAG 1302
Db 301 CGCAGCTGCAGATTAATGAAGATACAGGCTCTTGGACTGGCTTTCAACATTTGCCCATGAG 360

QY 1303 TCTGGACACAACTTTGGCAGATTTATGATGAGAGAGGGAACATGTGTAAAGATCCGAG 1362
Db 361 TCTGGACACAACTTTGGCAGATTTATGATGAGAGAGGGAACATGTGTAAAGATCCGAG 420

QY 1363 GGCAACATCATGTCCCTTACATTTGGCAGGAGCGCAATGGAGTCTTCTCTGGTCAACCCTGC 1422
Db 421 GGCAACATCATGTCCCTTACATTTGGCAGGAGCGCAATGGAGTCTTCTCTGGTCAACCCTGC 480

QY 1423 AGCCGCCAGTATCTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCCTTGTGATCAG 1482
Db 481 AGCCGCCAGTATCTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCCTTGTGATCAG 540

QY 1483 CCAAGGCTGTGAGGAGTAACAGTATCTTGGAGAAATTCAGGAGAGATATATATGATGCA 1542
Db 541 CCAAGGCTGTGAGGAGTAACAGTATCTTGGAGAAATTCAGGAGAGATATATATGATGCA 600
```

```

QY      1543 AACACACAGTGCAGTGGCGAGTTCCG 1567
DB      601 AACACACAGTGCAGTGGCGAGTTCCG 625

RESULT 7
BM845044
LOCUS   K-EST0123276 S12SNU216 Homo sapiens cDNA clone S12SNU216-97-E10 5',
DEFINITION mRNA sequence.
ACCESSION BM845044
VERSION   BM845044.1 GI:19201443
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
           1 (bases 1 to 623)
           Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
           Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
           Kim,Y.S.
TITLE     21C Frontier Korean EST Project 2001
JOURNAL   Unpublished (2002)
COMMENT   Contact: Kim YS
           Genome Research Center
           Korea Research Institute of Bioscience & Biotechnology
           52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
           Tel: +82-42-860-4470
           Fax: +82-42-860-4409
           Email: yongsung@mail.kribb.re.kr
           Plate: 97 row: E column: 10
           High quality sequence stop: 623.
           Location/Qualifiers
             1..623
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="S12SNU216-97-E10"
               /sex="F"
               /tissue_type="Lymph node"
               /cell_type="Epithelial"
               /cell_line="SNU-216"
               /lab_host="Top10F"
               /clone_lib="S12SNU216"
               /notes="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
               Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
               bacterial alkaline phosphatase (BAP) and then decapped
               with tobacco acid pyrophosphatase (TAP). The decapped
               intact mRNA was ligated with DNA-RNA linker including BcoR
               I site by treatment of T4 RNA ligase and the first strand
               cDNA was synthesized from oligo dt-selected mRNA by
               priming with dt-tailed vector. The dt-tailed vector was
               adjusted to have about 60nt. The cDNA vector was
               circularized with E. coli DNA ligase after digestion of
               EcoRI which site is also included in vector. An RNA strand
               converted to a DNA strand by Okayama-Berg method. The
               obtained cDNA vectors were used for transformation of
               competent cells E. coli Top10F by electroporation method.
               The cDNA libraries constructed by this method are
               full-length enriched cDNA library."

ORIGIN
Query Match 16.9%; Score 621.4; DB 4; Length 623;
Best Local Similarity 99.8%; Pred. No. 2.3e-161;
Matches 622; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      943 GTGCTCAGGATCTCAACATGGTATCTGCTTTTATTCAAGATGGAACAATAGGAGGAAC 1002
DB      1 GTGCTCAGGATCTCAACATGGTATCTGCTTTTATTCAAGATGGAACAATAGGAGGAAC 60
QY      1003 ATCAACATTGCAATTGTAGTCTGATCTTTCTAGAAAGATGAACAGCCAGGACCTGGTGATA 1062

```

```

DB      61 ATCAACATTGCAATTGTAGTCTGATCTTTCTAGAAAGATGAACAGCCAGGACCTGGTGATA 120
QY      1063 AGTCACCAACGACAGACACACCTTAAGTAGCTTTCTGCCAGTGGCAGCTCTGGATGATGGGG 1122
DB      121 AGTCACCAACGACAGACACACCTTAAGTAGCTTTCTGCCAGTGGCAGCTCTGGATGATGGGG 180
QY      1123 AAAGATGGGACCTGTCATGACACCGCCATCTTACTGACCTGGTCTGGATATATGTTCTCTGG 1182
DB      181 AAAGATGGGACCTGTCATGACACCGCCATCTTACTGACCTGGTCTGGATATATGTTCTCTGG 240
QY      1183 AAGAATGAGCCCTGTGACACTTTTGGGATTTGCACCCATAAGTGGAAATGTGTAGTAAATAT 1242
DB      241 AAGAATGAGCCCTGTGACACTTTTGGGATTTGCACCCATAAGTGGAAATGTGTAGTAAATAT 300
QY      1243 CGCAGTCGACCGATTAAATGAAGATACAGCTCTTGGACTGGCCTTCAACCATGCCCCATGAG 1302
DB      301 CGCAGTCGACCGATTAAATGAAGATACAGCTCTTGGACTGGCCTTCAACCATGCCCCATGAG 360
QY      1303 TCTGGACACAACTTTTGGGATGATTCATGATGAGGAAGGAAACATGTGTAAAAGTCCGAG 1362
DB      361 TCTGGACACAACTTTTGGGATGATTCATGATGAGGAAGGAAACATGTGTAAAAGTCCGAG 420
QY      1363 GGCACACATCATGTCCTCCCTACATTTGGCAGGACCAATGGAGTCTTCTCCTGGTCACCTGC 1422
DB      421 GGCACATCATGTCCTCCCTACATTTGGCAGGACCAATGGAGTCTTCTCCTGGTCACCTGC 480
QY      1423 AGCCGCCAGTATCTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCTGTCATGATCAG 1482
DB      481 AGCCGCCAGTATCTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCTGTCATGATCAG 540
QY      1483 CCAAGCCTGTGAAGGATACAAAGTATCTCTGAGAAATTCAGGAGGAATTTATATGATGCA 1542
DB      541 CCAAGCCTGTGAAGGATACAAAGTATCTCTGAGAAATTCAGGAGGAATTTATATGATGCA 600
QY      1543 AACACACAGTGCAGTGGCGAGTT 1565
DB      601 AACACACAGTGCAGTGGCGAGTT 623

RESULT 8
BM845406
LOCUS   K-EST0123712 S12SNU216 Homo sapiens cDNA clone S12SNU216-99-B04 5',
DEFINITION mRNA sequence.
ACCESSION BM845406
VERSION   BM845406.1 GI:19201805
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 537)
AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
           Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
           Kim,Y.S.
TITLE     21C Frontier Korean EST Project 2001
JOURNAL   Unpublished (2002)
COMMENT   Contact: Kim YS
           Genome Research Center
           Korea Research Institute of Bioscience & Biotechnology
           52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
           Tel: +82-42-860-4470
           Fax: +82-42-860-4409
           Email: yongsung@mail.kribb.re.kr
           Plate: 99 row: B column: 04
           High quality sequence stop: 537.
           Location/Qualifiers
             1..537
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="S12SNU216-99-B04"
               /sex="F"

FEATURES
source

```



```

/tissue_type="lymph node"
/cell_type="Epithelial"
/cell_line="SNU-216"
/lab_host="Top10P"
/clone_lib="S128N216"
/note="Organ: Stomach; Vector: pcns; Site 1: EcoRI,
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 50nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10p by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN
Query Match          14.6%; Score 535.4; DB 4; Length 537;
Best Local Similarity 99.8%; Pred. No. 2.2e-137;
Matches 536; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 943 GTGCTCAGATACCAACATGGTATCTGCTTTATTCAGAGTGAACATAGAGGAAC 1002
DB 1 GTGCTCAGATACCAACATGGTATCTGCTTTATTCAGAGTGAACATAGAGGAAC 60

QY 1003 ATCAACATGGAAATGTAGTCTGATCTTTCTAGAGATGAACAGCAGCTGGTGATA 1062
DB 61 ATCAACATGGAAATGTAGTCTGATCTTTCTAGAGATGAACAGCAGCTGGTGATA 120

QY 1063 AGTCACACGACAGACACCTTAAGTAGCTTCTGCGAGTGGCAGTCTGGATGGGG 1122
DB 121 AGTCACACGACAGACACCTTAAGTAGCTTCTGCGAGTGGCAGTCTGGATGGGG 180

QY 1123 AAAGATGGGACTCGTCATGACCGCATCTTACTGACTGTCTGGATATATGTCCTGG 1182
DB 181 AAAGATGGGACTCGTCATGACCGCATCTTACTGACTGTCTGGATATATGTCCTGG 240

QY 1183 AAGAAATGAGCCCTGTGACACTTTGGGATTTGACCCATTAAGTGGAAATGTAGTAATAT 1242
DB 241 AAGAAATGAGCCCTGTGACACTTTGGGATTTGACCCATTAAGTGGAAATGTAGTAATAT 300

QY 1243 CGCAGCTGCACGATTAATGAAGATACAGGTCTTTGGACTGSCCTTCACCATGGCCCATGAG 1302
DB 301 CGCAGCTGCACGATTAATGAAGATACAGGTCTTTGGACTGSCCTTCACCATGGCCCATGAG 360

QY 1303 TCTGGACACACTTTGGCATGATTCATGATGAGAGGGAACATGTGTAAGATGTCGAG 1362
DB 361 TCTGGACACACTTTGGCATGATTCATGATGAGAGGGAACATGTGTAAGATGTCGAG 420

QY 1363 GGCAACATCATGTGTCCTTACATTTGGCAGGACCAATGAGTCTTCTGCTGACCTGTC 1422
DB 421 GGCAACATCATGTGTCCTTACATTTGGCAGGACCAATGAGTCTTCTGCTGACCTGTC 480

QY 1423 AGCCGCGAGTATCTACAAATTTCTAAGCAGCGCTCAAGCTATCTGCTTGTCTGAT 1479
DB 481 AGCCGCGAGTATCTACAAATTTCTAAGCAGCGCTCAAGCTATCTGCTTGTCTGAT 537

```

```

RESULT 9
CD355731
LOCUS
DEFINITION
UI-N-FY0-cgo-o-23-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE:30355654 5', mRNA sequence.
ACCESSION
CD355731
VERSION
CD355731.1 GI:31148232
KEYWORDS
EST.
Mus musculus (house mouse)

```

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 728)
NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabp@rmail.nih.gov

REFERENCE

AUTHORS
TITLE
JOURNAL
COMMENT

Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project

Seq primer: pYX-5.

Location/Qualifiers
1..728
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30355654"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (TI phage resistant)"
/clone_lib="NIH_BMAP_FY0"
/note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Ronald, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

FEATURES

source

ORIGIN

Query Match 14.3%; Score 523.8; DB 6; Length 728;
Best Local Similarity 82.4%; Pred. No. 4.2e-134;
Matches 600; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 554 CACACCTCTCATGTGAACCTCGGAGAGCTGCCAAGGAGCTCGCCATCCACGTACTGT 613
DB 1 CACACCTAAACAGCAAACTGAAACAGATCTGCACAGGGGAGCTCCCTCCCATGTTCTGT 60

QY 614 ACAAGAGATCCACAGAGCCCTCATGCTCTTGGGGCAGTGAGTCTGTGAGCTCAAGA 673
DB 61 ACAAGAGATCTACAGAGCGCCAGGCTCTAGAGAAAAGAGTCTGTATGATCAGCAGA 120

QY 674 CATGGAGAGTGCACATCAACCCCTGACAGCGAGCTTTGGCTGGGAGCTGCCACAAA 733
DB 121 AGCGAGATCTGGCAAGAGCCACACCTGCACCATGACAACTTCCATCTTGGCCCTCACAAA 180

QY 734 AGCAGCATTTCTGTGGAAGAGCAAGAAATACATGCCCCAGCCCTCCCAAGAACCTCT 793
DB 181 AGCAGCATTTCTGTGGAAGAGCAAGAAATACATGCCCTCCCAAGAACCTCTCT 240

QY 794 TCATCTTCCAGATGAGTATAGTCTTGTTCACGCGATAAGCGTCTCTCTCTGAGGTC 853
DB 241 ACATCTTCCCTGATGAGTACAAGCTAGTTCCTCCGACACAGCGCTCGCTCTTGAATCCC 300

QY 854 ATAGAAATGAAGAACTGAACCTGGAGACCTTGGTGGTTCGACAAAGATGATGAAA 913
DB 301 ACAGAAATGAAGAGTTAATGTGGAGACCTGGTGGTGGTTGACAGAAAGATGATGAAA 360

QY 914 ACCATGGCCATGAAATATATACCACTTACCTGCTCAGGATCTCAACATGGTATCTGCTT 973
|||||
Db 361 GCCATGGCCATGAAACATTTACGACTTATGCTCTCACCATTCTCAACATGGTCTGCTCCT 420
|||||
QY 974 TATTCAAGATGCAACAAATAGGAGGAAACATCAACATTTGTAAGTCTGATTTCTTC 1033
|||||
Db 421 TATTCAAGATGCAACAAATAGGAGGAAACATCAACATTTGTAAGTCTGATTTCTTC 480
|||||
QY 1034 TAGAAGATGAACAGCCAGGACTGGTATAGTCAACACGAGACCAACACCTTAAGTAGCT 1093
|||||
Db 481 TGAAGATGAACAGCCAGGCTGCGATTAGTCAACATGACAGACCAACAGCTTAACAGCT 540
|||||
QY 1094 TCTGCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTGCTCATGACCGCCATCT 1153
|||||
Db 541 TCTGCCAGTGGCAGTCTGGATTGATGGGAAAGATGGGACTGCTCATGACCGCCATCT 600
|||||
QY 1154 TACTGACTGCTGGATATATATGTTCTCTGGAAGAAATGAGCCCTGTGACACTTTGGGATTG 1213
|||||
Db 601 TACTGACTGCTGGATATATGTTCTCTGGAAGAAATGAGCCCTGTGACACTTTGGGATTG 660
|||||
QY 1214 CACCCATAAGTGAATGTGTAGTAAATATGCGAGCTGCGAGTAAATGAAGATACAGGTC 1273
|||||
Db 661 CACCCATAAGCGGATGTGCAGTAAATGACGAGCTGCGAGTCAATGNAGACTCAGGAC 720
|||||
QY 1274 TTGGACTG 1281
|||||
Db 721 TTGGACTG 728
|||||

RESULT 10
CD636691/c
LOCUS
DEFINITION 523 bp mRNA linear EST 12-JAN-2004
ACCESSION
CD636691
VERSION
CD636691.1 GI:40284958
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 523)
Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
Genomics 84 (1), 205-210 (2004)
JOURNAL
COMMENT
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
Location/Qualifiers
1..523
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/notes="Vector: pDrive Cloning Vector"

ORIGIN
Query Match 13.9%; Score 511.2; DB 6; Length 523;
Best Local Similarity 99.4%; Pred. No. 1.2e-130;
Matches 513; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 177 ATATGACCTGGTCTCTGCTTACGAGTTGACACAGGGGGGATTTAGTGTCCCATGAAT 236
|||||
Db 516 ATATGACCTGGTCTCTGCTTACGAGTTGACACAGGGGGGATTTAGTGTCCCATGAAT 457
|||||
QY 237 CATGCAACCATCAGCGCGGAGAGAGCAGTGGCCGTGTCCGAGTTGAGTCTCTTCACTT 296
|||||
Db 456 CATGCAACCATCAGCGCGGAGAGAGAGCAGTGGCCGTGTCCGAGTTGAGTCTCTTCACTT 397
|||||
QY 297 TCGGCTGAAAGGCTCCAGGACGACTTCCACGTGGATCTGAGGACTTCCAGCAGCCTAGT 356
|||||

Db 396 TCGGCTGAAAGGCTCCAGGACGACTTCCACATGGATCTGAGGACTTCCAGCAGCCTAGT 337
|||||
QY 357 GGCTCTCTGCTTTATTGTGCAGACGTTGGGAAAGACAGGCACTTAAGTCTGTGTGAGACTTTT 416
|||||
Db 336 GGCTCTCTGCTTTATTGTGCAGACGTTGGGAAAGACAGGCACTTAAGTCTGTGTGAGACTTTT 277
|||||
QY 417 ACGCCAGAGAGACTTCTGTTTCTATCAAGGCTCTTTGGCATCAACAGAAAATCTCTCAGT 476
|||||
Db 276 ACGCCAGAGAGACTTCTGTTTCTATCAAGGCTCTTTGGCATCAACAGAAAATCTCTCAGT 217
|||||
QY 477 GCGCCCTTCAACCTGCGAAGGCTTGTGAGGATGATACGACAGAGAGGAGAGGACTTACTT 536
|||||
Db 216 GCGCCCTTCAACCTGCGAAGGCTTGTGAGGATGATACGACAGAGAGGAGAGGACTTACTT 157
|||||
QY 537 CTTAAGGCCACTTCTCTTCAACCTCTCATGGAACCTCGGCAGAGCTGCCCAAGGCAAGCTC 596
|||||
Db 156 CTTAAGGCCACTTCTCTTCAACCTCTCATGGAACCTCGGCAGAGCTGCCCAAGGCAAGCTC 97
|||||
QY 597 GGCATCCAGCTACTGTACAAGAGATCCAGAGCCCCATGTCTCTGGGGCCAGTGAGGT 656
|||||
Db 96 GGCATCCAGCTACTGTACAAGAGATCCAGAGCCCCATGTCTCTGGGGCCAGTGAGGT 37
|||||
QY 657 CTGTGTGACTCAAGGACATGGAGCTGGCAGCATCA 692
|||||
Db 36 CTGTGTGACTCAAGGACATGGAGCTGGCAGCATCA 1
|||||

RESULT 11
CD636687/c
LOCUS
DEFINITION 523 bp mRNA linear EST 12-JAN-2004
ACCESSION
CD636687
VERSION
CD636687.1 GI:40284954
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 523)
Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
Genomics 84 (1), 205-210 (2004)
JOURNAL
COMMENT
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
Location/Qualifiers
1..523
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN
Query Match 13.9%; Score 509.6; DB 6; Length 523;
Best Local Similarity 99.2%; Pred. No. 3.4e-130;
Matches 512; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 177 ATATGACCTGGTCTCTGCTTACGAGTTGACACAGGGGGGATTTAGTGTCCCATGAAT 236
|||||
Db 516 ATATGACCTGGTCTCTGCTTACGAGTTGACACAGGGGGGATTTAGTGTCCCATGAAT 457
|||||
QY 237 CATGCAACCATCAGCGCGGAGAGAGCAGTGGCCGTGTCCGAGTTGAGTCTCTTCACTT 296
|||||
Db 456 CATGCAACCATCAGCGCGGAGAGAGCAGTGGCCGTGTCCGAGTTGAGTCTCTTCACTT 397
|||||
QY 297 TCGGCTGAAAGGCTCCAGGACGACTTCCACGTGGATCTGAGGACTTCCAGCAGCCTAGT 356
|||||
Db 396 TCGGCTGAAAGGCTCCAGGACGACTTCCACATGGATCTGAGGACTTCCAGCAGCCTAGT 337
|||||

```

QY 357 GGCTCTGGCTTTATTGTGCAGACGTTGGGAAGACAGGCACTAAGTCTGTGCAGACTTT 416
|
|
|
Db 336 GGCTCTCTGGCTTTATTGTGCAGACGTTGGGAAGACAGGCACTAAGTCTGTGCAGACTTT 277
|
|
|
QY 417 ACCGCCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGCATCACACAGAACTCCTCAGT 476
|
|
|
Db 276 ACCGCCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGCATCACACAGAACTCCTCAGT 217
|
|
|
QY 477 GGCCCTTTCAACCTGCCAAGGCTTGTGAGGATGATAGAAACAGAGGCGAGATTACTT 536
|
|
|
Db 216 GGCCCTTTCAACCTGCCAAGGCTTGTGAGGATGATAGAAACAGAGGCGAGATTACTT 157
|
|
|
QY 537 CCTAAGGCCACTTCTCTTACACCTCTCATGGAACCTCGGAGAGCTGCCAAGGCGAGCTC 596
|
|
|
Db 156 CCTAAGGCCACTTCTCTTACACCTCTCGTGAACCTCGGAGAGCTGCCAAGGCGAGCTC 97
|
|
|
QY 597 GCCATCCCAAGTACTGTACAAAGATCCACAGAGCCCCCATGCTCTGGGGCCAGTGAGGT 656
|
|
|
Db 96 GCCATCCCAAGTACTGTACAAAGATCCACAGAGCCCCCATGCTCTGGGGCCAGTGAGGT 37
|
|
|
QY 657 CCTGTGACCTCAAGACATGGAGCTGGCACATCA 692
|
|
|
Db 36 CCTGGTGAACCTCAAGGACATGGAGCTGGCACATCA 1
|
|
|

```

```

RESULT 12
CD636689/c
LOCUS
DEFINITION 523 bp mRNA linear EST 12-JAN-2004
ACCESSION
VERSION
CD636689.1 GI:40284956
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE
AUTHORS Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
TITLE Circular rapid amplification of cDNA ends for high-throughput
JOURNAL extension cloning of partial genes
COMMENT Genomics 84 (1), 205-210 (2004)
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com

```

```

FEATURES
source
1..523
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

```

```

ORIGIN
Query Match 13.8%; Score 508; DB 6; Length 523;
Best Local Similarity 99.0%; Pred. No. 9.5e-130;
Matches 511; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 177 ATATGACCTGGTCTCTGCTACAGGTTGACACAGGGGCGGATTAAGTGTCCCATGAAAT 236
|
|
|
Db 516 ATATGACCTGGTCTCTGCTACAGGTTGACACAGGGGCGGATTAAGTGTCCCATGAAAT 457
|
|
|
QY 237 CATGACCAATCATCGCGCGGAGAGAGCAGTGGCCGTGTCGAGGTTGAGTCTCTTCACCT 296
|
|
|
Db 456 CATGACCAATCATCGCGCGGAGAGAGCAGTGGCCGTGTCGAGGTTGAGTCTCTTCACCT 397
|
|
|
QY 297 TCGGTGTAAGGCTCCAGGCAAGTCTTCCAGTGGATCTGAGGATTTCCAGCAGCCTAGT 356
|
|
|
Db 396 TCGGTGTAAGGCGCCAGGCAAGTCTTCCAGTGGATCTGAGGATTTCCAGCAGCCTAGT 337
|
|
|
QY 357 GGCTCTCTGGCTTTATTGTGCAGACGTTGGGAAGACAGGCACTAAGTCTGTGCAGACTTT 416
|
|
|

```

```

Db 336 GGCTCTCTGGCTTTATTGTGCAGACGTTGGGAAGACAGGCACTAAGTCTGTGCAGACTTT 277
|
|
|
QY 417 ACCGCCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGCATCACACAGAACTCCTCAGT 476
|
|
|
Db 276 ACCGCCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGCATCACACAGAACTCCTCAGT 217
|
|
|
QY 477 GGCCCTTTCAACCTGCCAAGGCTTGTGAGGATGATAGAAACAGAGGCGAGATTACTT 536
|
|
|
Db 216 GGCCCTTTCAACCTGCCAAGGCTTGTGAGGATGATAGAAACAGAGGCGAGATTACTT 157
|
|
|
QY 537 CCTAAGGCCACTTCTCTTACACCTCTCATGGAACCTCGGAGAGCTGCCAAGGCGAGCTC 596
|
|
|
Db 156 CCTAAGGCCACTTCTCTTACACCTCTCATGGAACCTCGGAGAGCTGCCAAGGCGAGCTC 97
|
|
|
QY 597 GCCATCCCAAGTACTGTACAAAGATCCACAGAGCCCCCATGCTCTGGGGCCAGTGAGGT 656
|
|
|
Db 96 GCCATCCCAAGTACTGTACAAAGATCCACAGAGCCCCCATGCTCTGGGGCCAGTGAGGT 37
|
|
|
QY 657 CCTGTGACCTCAAGGACATGGAGCTGGCACATCA 692
|
|
|
Db 36 CCTGGTGAACCTCAAGGACATGGAGCTGGCACATCA 1
|
|
|

```

```

RESULT 13
CD636693/c
LOCUS
DEFINITION 522 bp mRNA linear EST 12-JAN-2004
ACCESSION
VERSION
CD636693.1 GI:40284960
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE
AUTHORS Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
TITLE Circular rapid amplification of cDNA ends for high-throughput
JOURNAL extension cloning of partial genes
COMMENT Genomics 84 (1), 205-210 (2004)
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com

```

```

FEATURES
source
1..522
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

```

```

ORIGIN
Query Match 13.8%; Score 507; DB 6; Length 522;
Best Local Similarity 99.0%; Pred. No. 1.8e-129;
Matches 510; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 177 ATATGACCTGGTCTCTGCTACAGGTTGACACAGGGGCGGATTAAGTGTCCCATGAAAT 236
|
|
|
Db 515 ATATGACCTGGTCTCTGCTACAGGTTGACACAGGGGCGGATTAAGTGTCCCATGAAAT 456
|
|
|
QY 237 CATGACCAATCATCGCGCGGAGAGAGCAGTGGCCGTGTCGAGGTTGAGTCTCTTCACCT 296
|
|
|
Db 455 CATGACCAATCATCGCGCGGAGAGAGCAGTGGCCGTGTCGAGGTTGAGTCTCTTCACCT 396
|
|
|
QY 297 TCGGTGTAAGGCTCCAGGCAAGTCTTCCAGTGGATCTGAGGATTTCCAGCAGCCTAGT 356
|
|
|
Db 395 TCGGTGTAAGGCGCCAGGCAAGTCTTCCAGTGGATCTGAGGATTTCCAGCAGCCTAGT 336
|
|
|
QY 357 GGCTCTCTGGCTTTATTGTGCAGACGTTGGGAAGACAGGCACTAAGTCTGTGCAGACTTT 416
|
|
|
Db 335 GGCTCTCTGGCTTTATTGTGCAGACGTTGGGAAGACAGGCACTAAGTCTGTGCAGACTTT 276
|
|
|

```

```

QY 417 ACCGCCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGATCAACAGAACTCCTCAGT 476
|
|
|
Db 275 ACCGCCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGATCAACAGAACTCCTCAGT 216
|
|
|
QY 477 GGCCTTTTCAACTGCGCAAGGCTTGTGAGGATGATACGAAACAGAGAGGAGCAGATTACTT 536
|
|
|
Db 215 GGCCTTTTCAACTGCGCAAGGCTTGTGAGGATGATACGAAACAGAGAGGAGCAGATTACTT 156
|
|
|
QY 537 CCTAAGGCCACTTCTTTCACACTCTCATGGAACATCGGCAGAGCTGCCCAAGGCAGCTC 596
|
|
|
Db 155 CCTAAGGCCACTTCTTTCACACTCTCATGGAACATCGGCAGAGCTGCCCAAGGCAGCTC 96
|
|
|
QY 597 GCATCCCAAGTACTGTACAGAGATCCACAGAGCCCATGCTCTGGGGCCAGTGAGGT 656
|
|
|
Db 95 GCATCCCAAGTACTGTACAGAGATCCACAGAGCCCATGCTCTGGGGCCAGTGAGGT 36
|
|
|
QY 657 CTGGTGACCTCAAGGACATGGGAGCTGGGCACATC 691
|
|
|
Db 35 CTGGTGACCTCAAGGACATGGGAGCTGGGCAATC 1
|
|
|
RESULT 14
CD636688
LOCUS CD636688 523 bp mRNA linear EST 12-JAN-2004
DEFINITION S6008668J1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD636688
VERSION CD636688.1 GI:40284955
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
COMMENT Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
FEATURES
source
Location/Qualifiers
1..523
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/notes="Vector: pDrive Cloning Vector"
ORIGIN
Query Match 13.6%; Score 498.6; DB 6; Length 523;
Best Local Similarity 99.0%; Pred. No. 4e-127;
Matches 512; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 177 ATATGACCTGGTCTCTGCTTACGAGTTGACACAGGGCGGATAGCTGTCCCATGAAT 236
|
|
|
Db 8 ATATGACCTGGTCTCTGCTTACGAGTTGACACAGGGCGGATAGCTGTCCCATGAAT 67
|
|
|
QY 237 CATGCAACATCAGCGCGGAGAGAGAGAGTGGCGGTGTCGAGTTGAGTCTCTTCACT 296
|
|
|
Db 68 CATGCAACATCAGCGCGGAGAGAGAGAGTGGCGGTGTCGAGTTGAGTCTCTTCACT 127
|
|
|
QY 297 TCGGCTGAAGGCTCCAGGACACACTCCACGTGGATCTGAGACCTCCAGAGCCTACT 356
|
|
|
Db 128 TCGGCTGAAGGCTCCAGGACACACTCCACGTGGATCTGAGACCTCCAGAGCCTACT 187
|
|
|
QY 357 GGCCTCTGGCTTATTTGTGACAGCTTGGGAAGACAGGCACTAAGTCTGTGACAGCTTT 416
|
|
|
Db 188 GGCCTCTGGCTTATTTGTGACAGCTTGGGAAGACAGGCACTAAGTCTGTGACAGCTTT 247
|
|
|
QY 417 ACCGCCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGATCAACAGAACTCCTCAGT 476

```

```

Db 248 ACCGCCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGATCAACAGAACTCCTCAGT 307
|
|
|
QY 477 GGCCTTTTCAACTGCGCAAGGCTTGTGAGGATGATACGAAACAGAGAGGAGCAGATTACTT 536
|
|
|
Db 308 GGCCTTTTCAACTGCGCAAGGCTTGTGAGGATGATACGAAACAGAGAGGAGCAGATTACTT 367
|
|
|
QY 537 CCTAAGGCCACTTCTTTCACACTCTCATGGAACATCGGCAGAGCTGCCCAAGGCAGCTC 596
|
|
|
Db 368 CCTAAGGCCACTTCTTTCACACTCTCATGGAACATCGGCAGAGCTGCCCAAGGCAGCTC 427
|
|
|
QY 597 GCATCCCAAGTACTGTACAGAGATCCACAGAGCCCATGCTCTGGGGCCAGTGAGGT 656
|
|
|
Db 428 GCATCCCAAGTACTGTACAGAGATCCACAGAGCCCATGCTCTGGGGCCAGTGAGGT 486
|
|
|
QY 657 CTGGTGACCTCAAGGACATGGGAGCTGGGCACATCAA 693
|
|
|
Db 487 CTGGTGACCTCAAGGACATGGGAGCTGGGCACATCAA 523
|
|
|
RESULT 15
CF170272
LOCUS CF170272 630 bp mRNA linear EST 25-JUL-2003
DEFINITION B0825B02-5 NIA Mouse Newborn Kidney cDNA Library (Long 1) Mus
musculus cDNA clone NIA:B0825B02 IMAGE:30469837 5', mRNA sequence.
ACCESSION CF170272
VERSION CF170272.1 GI:33279821
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method
JOURNAL Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE 21429098
PUBMED 11544199
COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: B0825 row: B column: 02
Seq primer: M13 Reverse
High quality sequence stop: 630
POLYA=No.
FEATURES
Location/Qualifiers
1..630
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="CS7BL/6J"
/db_xref="niaEST:B0825B02-5"
/db_xref="taxon:10090"
/clone="NIA:B0825B02 IMAGE:30469837"
/dev_stage="Newborn Kidney"
/lab_host="DH10B"
/clone_lib="NIA Mouse Newborn Kidney cDNA Library (Long
1)"
/notes="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
Site 2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
In brief, double-stranded cDNAs were synthesized with an
Oligo(dT) primer (Invitrogen):
5'-pGACTAGTTCTAGATCGGAGCGGCCCTTTT-3' from
26 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Lone-linker LL-SalI, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using

```

Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with Sali and NotI enzymes and cloned into Sali/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.0 kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match	13.5%	Score 497.8	DB 7	Length 630
Best Local Similarity	87.0%	Pred. No. 7.2e-127		
Matches 547	Conservative 0	Mismatches 82	Indels 0	Gaps 0
Qy	1600	AAAAAGGACATCTGTAAAGCCCTGTGGTCCATCGTATTGGAAGAAATGTGAGACTAAA	1659	
Db	2	AGAAAGGACATCTGTAAAGCCCTGTGGTCCATCGGATTGGAAGAAATGTGAGACCAAG	61	
Qy	1660	TTTATGCCAGCAGCAGAGGACCAATTTGTGGGCATGACATGTGTGCGGGAGGACAG	1719	
Db	62	TTTATGCCAGCAGCAGAGGACCAATTTGTGGGCATGACATGTGTGCGGGAGGACAG	121	
Qy	1720	TGTGTGAATATGTTGATCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAG	1779	
Db	122	TGTGTCAAGTACGGTGTGAAGGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAG	181	
Qy	1780	TCTTGTGTCCTCCATGCTCCAGGACCTTGGAGGGGAGTATCTCATAGGAGTCGCTCTGC	1839	
Db	182	CCCTGTGTCCTCCATGCTCCAGGACCTTGGAGGGGAGTATCTCATAGGAGTCGCTCTGC	241	
Qy	1840	ACCAACCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAG	1899	
Db	242	ACCAATCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAG	301	
Qy	1900	CTCTGCAACAGTCAAGAAATGTCCTGGGACAGTGTTCACCTTCCGTCGCTCAGTGTGCC	1959	
Db	302	CTTTGCAACAGTCAAGAAATGTCCTGGGACAGTGTTCACCTTCCGTCGCTCAGTGTGCC	361	
Qy	1960	GAGCACACAGCAGACGATTCAGAGGGCGGCACTACAAGTGGAGCCCTTACACCCAGTTG	2019	
Db	362	GAGTATACAGCAGACGCTTCGAGGGTGGCTCTACAAGTGGAGCCCTTACACCCAGTTG	421	
Qy	2020	GAAGATCAGGACTTATGCAAACTTACTGTATCGCAGAGGATTTGATTTCTTTCT	2079	
Db	422	GAAGATCAGGACTTATGCAAACTTACTGTATCGCAGAGGATTTGATTTCTTTCT	481	
Qy	2080	TTGTCAAAATAAGTCAAGATGGACTCCATGCTCGAGGATAGCCGTAATGTTGTATA	2139	
Db	482	TTGTCAAAATAAGTCAAGATGGACTCCATGCTCGAGGATAGCCGTAATGTTGTATA	541	
Qy	2140	GATGGGATGTGAGAGAGTGGATGTGACAAATGTCCTTGGATCTGATGCTGTTGAAGAC	2199	
Db	542	GATGGGATGTGAGAGAGTGGCTGTGACAAATGTCCTTGGATCTGATGCTGTTGAAGAC	601	
Qy	2200	GTCGTGGGGTGTAAACGGGAATTAATC	2228	
Db	602	TCCTGTGGGGTGTGCAAGGGGAATTAATC	630	

Search completed: June 14, 2005, 00:20:10
Job time : 10417 secs